

第1部門第1区分

(43)公表日 平成6年(1994)3月24日

| | | | |
|--------------|-----------|--------|-----|
| (51)Int.Cl.* | 識別記号 | 序内整理番号 | F I |
| C 12 N 15/77 | Z NA | | |
| A 61 K 39/00 | A 9284-4C | | |
| 39/395 | A 9284-4C | | |
| | H 9284-4C | | |
| C 12 N 1/21 | 7236-4B | | |

審査請求 未請求 予備審査請求 未請求(全32頁) 最終頁に続く

(21)出願番号 特願平5-503324
 (86) (22)出願日 平成4年(1992)7月29日
 (85)翻訳文提出日 平成5年(1993)3月30日
 (86)国際出願番号 PCT/FR92/00744
 (87)国際公開番号 WO93/03158
 (87)国際公開日 平成5年(1993)2月18日
 (31)優先権主張番号 91/09652
 (32)優先日 1991年7月30日
 (33)優先権主張国 フランス(FR)
 (31)優先権主張番号 91/09870
 (32)優先日 1991年8月2日
 (33)優先権主張国 フランス(FR)

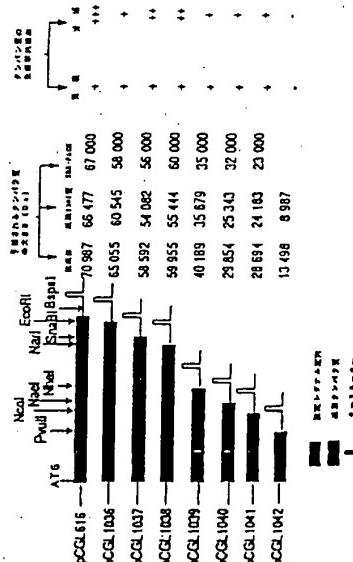
(71)出願人 オルサン
 フランス国パリ、リュ、パリュ、16
 (72)発明者 ジョリフ、グワンナエル
 フランス国パリ、リュ、トリュフォー、48
 (72)発明者 ギヨンバルシュ、アルメル
 フランス国ラーイ、レ、ローズ、アブニ
 ュ、フルーケ、21
 (72)発明者 レラノ、ビュリフィカション
 フランス国フォントネ、オ、ローズ、リ
 ュ、ピエールレ、12
 (74)代理人 弁理士 佐藤一雄(外2名)

最終頁に続く

(54)【発明の名称】特にコリネバクテリア中で用いることのできる蛋白質の発現および分泌系

(57)【要約】

コリネバクテリアによるアミノ酸、ポリペプチドまたは蛋白質の発現および分泌のための系である。この系は、そのアミノ酸、ポリペプチドまたは蛋白質をコードする配列が染色体またはプラスミドDNAの領域内に置かれ、その配列が、蛋白質PS1またはPS2のシグナル配列をコードする配列の少なくとも1つの部分(この部分はコリネバクテリア菌株へこの系が取り込まれた時、翻訳後のその蛋白質の分泌を保証する)とともに5'末端に向かって転写されることを特徴とするものである。



請求の範囲

1. コリネバクテリア菌株による所定のアミノ酸、ポリペプチドまたは蛋白質の発現および分泌のための系であって、上記アミノ酸、ポリペプチドまたは蛋白質をコードする配列が、染色体またはプラスミドDNAの領域内に位置しており、この領域内では該配列が、蛋白質PS1またはPS2のシグナル配列をコードする配列の少なくとも1つの部分とともに5'末端に向かって転写され、該部分は、該系がコリネバクテリア菌株に取り込まれた時、翻訳後の上記蛋白質の分泌を保証するものであることを特徴とする、系。

2. 次の構成から成るコリネバクテリアの発現および分泌系であって

- コリネバクテリア菌株と、
- 上記コリネバクテリア菌株中の発現のための第一複数性DNA配列、アミノ酸、ポリペプチドおよび／または蛋白質をコードする第二DNA配列、および上記第一および第二DNA配列の間に挿入された第三DNA配列を含む分泌カセットとを含んでなり、上記第三DNA配列は、上記コリネバクテリア菌株により上記アミノ酸、ポリペプチドおよび／または蛋白質の分泌を保証するPS1またはPS2から選ばれた蛋白質の要素をコードするものである、系。

項1～9のいずれか一項記載の系。

11. 所定のアミノ酸、ポリペプチドまたは蛋白質の発現および分泌が、温度、培養液および糖の性質によって制御される、請求項1～10のいずれか一項記載の系。

12. コード配列が1つ以上の反復アミノ酸のポリマーをコードする配列である、請求項1～11のいずれか一項記載の系。

13. 反復配列における反復単位が、COO甘味部において正または負に荷電したアミノ酸を含んでなる、請求項1～12のいずれか一項記載の系。

14. ポリペプチドのイオン特性がその準離を可能にする、請求項1～13のいずれか一項記載の系。

15. 荷電したアミノ酸が、特定のプロテアーゼによりペプチドの切断をそのレベルで可能にする、請求項1～14のいずれか一項記載の系。

16. 荷電したアミノ酸が特定のカルボキシペプチダーゼにより除去できる、請求項1～15のいずれか一項記載の系。

17. マーカー遺伝子がcsp1遺伝子である、請求項1～16のいずれか一項記載の系。

18. コード配列が、第17図に対応するgdhAのすべてまたは一部を含んでなる、請求項1～17のいずれか一項記載の系。

特表平6-502548 (2)

3. コリネバクテリアの菌株が*Brevibacterium*属に属するものである、請求項1または2に記載の発現および分泌系。

4. 発現のための第一複数性DNA配列がプロモーターおよびリボソーム結合部位を含んでなる、請求項1～3のいずれか一項記載の発現および分泌系。

5. 分泌カセットがコリネバクテリア菌株中で複数する複数起承を含む自律的に複製するプラスミドにより保持されている、請求項1～4のいずれか一項記載の発現および分泌系。

6. 分泌カセットがコリネバクテリア菌株の染色体へのその組込みを保証するDNAの要素を含む、請求項1～5のいずれか一項記載の発現および分泌系。

7. 上記第三DNA配列がPS1またはPS2のシグナル配列のすべてまたは一部を含んでなる、請求項1～6のいずれか一項記載の発現および分泌系。

8. コード配列の端部の翻訳停止配列、翻訳停止配列およびマーカー遺伝子をさらに含んでなる、請求項1～7のいずれか一項記載の発現および分泌系。

9. 所定のアミノ酸、ポリペプチドまたは蛋白質をコードする配列が、同一相でcsp1またはcsp2遺伝子に挿入されている、請求項1～8のいずれか一項記載の系。

10. PS1またはPS2配列が切形である、請求

19. 発現のための調節性DNA配列が、csp1、csp2またはgdhAの発現要素から選ばれるものである、請求項1～18のいずれか一項記載の系。

20. 発現が塩、代謝産物および糖の濃度に依存する、請求項1～19のいずれか一項記載の系。

21. コード配列の発現前において、プロモーターが、csp1、csp2またはgdhAプロモーターから選ばれるものである、請求項1～20のいずれか一項記載の系。

22. マーカー遺伝子がlacZ遺伝子である、請求項1～21のいずれか一項記載の系。

23. 請求項1～22のいずれか一項記載の発現および分泌系を使用して得られる、バクテリア菌株。

24. 菌株がコリネバクテリアである、請求項23に記載の菌株。

25. 菌株がブレビバクテリアである、請求項24に記載の菌株。

26. 菌株が*Brevibacterium lactofermentus*である、請求項25に記載のコリネバクテリア菌株。

27. 所定の蛋白質が、その固定作用を有するPS1またはPS2部位により壁上に固定されている、請求項23～26のいずれか一項記載のコリネバクテリア菌株。

28. 菌株が、その壁上に固定されたPS1または

特にコリネバクテリア中で用いることのできる蛋白質の発現および分泌系

本発明は、特にコリネバクテリア中で使用することのできる蛋白質の発現および分泌系、この系を使用する方法、およびこれら発現系に関連した新規な蛋白質に関するもの。

コリネバクテリアとは、さまざまな菌株によって表示される不規則な形態の一層のグラム陽性バクテリアのことである。

グラム陽性細胞は、外部培地への蛋白質の分泌を容易にする簡単な構造を有するという事実にもかかわらず、コリネバクテリアによる蛋白質の分泌は、今日迄あまり広く研究されていない。毒性^{*}の溶原性ファージ(Smith 1980; J. Bacteriol., 141, 1142頁; Smith等, 1980; J. Bacteriol., 141, 184 頁; Greenfield 等, 1983; PNAS USA 80, 8853頁)に感染した Corynebacterium diphtheriae の特定の菌株によって分泌されたジフェリニア素および Corynebacterium glutamicum (V. Liebl等、A.S. Slinkay, 1988; バクテリアの遺伝子学および生物工学、2巻、883-884 頁)によるDNAアーゼの分泌に伴う遺伝子のヌクレオチド配列の研究が報告されているに

PS2の抗原エピトープを有する、請求項23~26のいずれか一項記載のコリネバクテリア菌株。

29. アミノ酸、ポリペプチドまたは蛋白質を生成する方法であって、請求項23~28のいずれか一項記載のコリネバクテリア菌株を培養液中で培養し、第二DNAA配列が上記アミノ酸、ポリペプチドまたは蛋白質をコードし、培養後、上記生成物を、培養液および/またはバクテリア細胞から任意に分離することを含んでなる、方法。

30. PS1またはPS2と融合または別な方法で結合した所定の蛋白質が、表面活性剤を使用してバクテリア細胞から分離される、請求項29に記載の方法。

31. PS1またはPS2配列のすべてまたは一部を含む、蛋白質。

32. PS1またはPS2の抗原部位を含む、蛋白質。

33. 抗原要素として、請求項31または32記載の蛋白質。

34. PS1またはPS2に対する、抗体。

すぎない。

米国特許第4,965,197号は、上記DNAアーゼの下で Corynebacterium に用いることのできる発現および分泌系について述べているが、この場合の蛋白質は主成分ではなく、またこれら条件の下では、対応する分泌系はあまり重要ではないように思われる。

従って、本発明は、特定のコリネバクテリアの培養液上清ろ中に高い割合で存在する二つの蛋白質の分泌のための要素を含むコリネバクテリア型のバクテリアにおける発現および分泌系に関する。

本発明は、特にコリネバクテリア菌株により所定のアミノ酸、ポリペプチドまたは蛋白質の発現および分泌のための系であって、上記アミノ酸、ポリペプチドまたは蛋白質をコードする配列が、染色体またはプラスミドDNAの領域内に位置しており、この領域内では上記配列が、蛋白質PS1またはPS2のシグナル配列をコードする配列の少なくとも1つの部分とともに5'端部に向かって転写され、上記部分は、系が上記コリネバクテリア菌株に取り込まれた時、翻訳後の上記アミノ酸、ポリペプチドまたは蛋白質の分泌が保証するものであることを特徴とする系に関する。

さらに、本発明は、コリネバクテリアの発現および分泌系であって、

- コリネバクテリア菌株と、

- 上記コリネバクテリア菌株中の発現のための第一機能性DNA配列、アミノ酸、ポリペプチドおよび/または蛋白質をコードする第二DNA配列、および上記第一および第二DNA配列間に挿入された第三DNA配列を含む分泌カセットとを含んでなり、上記第三DNA配列は、上記コリネバクテリア菌株により上記アミノ酸、ポリペプチドおよび/または蛋白質の分泌を保証するPS1またはPS2から選ばれた蛋白質の要素をコードするもの、である系に関する。

まず、本発明の構成において、「コリネバクテリア」とは Corynebacterium 属の菌株のみならず Bravibacterium のような同属のバクテリアの菌株をも指すものと理解すべきである。

本発明の発現系はコリネバクテリア中で自動的に複製するプラスミド中に存在し、この場合、プラスミドは、例えば Corynebacterium の菌株中で複製する複製型、即ち複製型起源PBL1を含むが、また上記発現系は、染色体組込み用に特別に設計された複製不能なプラスミドに保持することもでき、この場合、プラスミドは染色体組換えおよび組込みを可逆にする要素を含んでなる。この組込みの場合、発現系は最終的に上記バクテリアの染色体中に存在する。

特に、染色体組込みの場合、PS1をコードする遺伝子CSP1またはPS2をコードする遺伝子CSP2へ

められる蛋白質 P S 1 および／または P S 2 の一方のプロモーターを使用することも可能である。

また、発現要素は遺伝子の下流領域の発現の調整を保証する DNA 配列を含むこともできる。

良好な発現を保証する要素として、コード配列の末端に 1 つ以上の停止コドンの形の組合停止要素、または転写停止要素を配置することが可能である。

分離を保証する要素としては、上記したように、分離特性を変更または喪失することなしに、蛋白質 P S 1 または P S 2 の一方のシグナル配列のすべてまたは一部並びにこれら配列と等価の配列を挙げることができる。

最終的には、点突然変異のような公知の技術を用いて、同様の分離特性を保持しながら分離配列を僅かに変更することが可能であり、従って、本発明はこれら等価の配列をも含むものである。

結論として、本発明による発現系は他の要素、特に転写ターミネーター、例えば蛋白質 P S 1 および／または P S 2 用、もしくは g d h A 用のターミネーターのような要素を含んでいてもよい。

場合によっては、発現および分離配列に蛋白質 P S 1 のすべてまたは一部を導入して、これらの条件の下で分離および発現レベルを改善することのできる融合蛋白質を得ることも有利である。

本発明による発現系は、例えば、E. coli 中で作

の異種 DNA 配列の挿入は、対応する菌株の成長に影響を及ぼさないことが実証されている。これらの条件の下では、挿入したコード配列の発現による生成物を発見／分離させるために、C S P 1 または C S P 2 の相の中でアミノ酸、ポリペプチドまたは蛋白質をコードする配列を一体化することが可能である。

コリネバクテリア菌株中の発現のための機能性 DNA 配列としては、相同発現用要素および異種発現用要素の両方を挙げができる。即ち、これら要素は宿主のバクテリア中にすでに存在している要素、またはこれとは異なり、異なるバクテリアから誘導される要素でもよい。

これら発現要素は、プロモーターおよびリボソーム結合部位を本質的に含んでいるが、他の要素、特に発現を調整するタイプの要素であることも可能である。

コリネバクテリア中で使用することのできる発現要素としては、強力なプロモーターである P lac プロモータ、I P T G により誘導され且つ E. coli のようなコリネバクテリア中で作用することが解っている t r p / lac ハイブリッドが特に使用される。しかしながら、他のプロモータ、例えば下記のプロモータ、またはコリネバクテリアの構造遺伝子発現のための他の要素、例えば g d h A プロモータを使用することが可能である。また、例えば、発現要素、特に本発明の範囲内にあると認

用する上記のような複数起源をコリネバクテリアと異なるバクテリア中で生成する異種要素、また Corynebacter lus への転移を容易にする標識遺伝子のような他の要素を含んでいてもよい。

もちろん、マーカー遺伝子は、コリネバクテリア中で作用する限りにおいて、様々な形のものであってよく、耐性のような正または負の選択用の遺伝子であってよい。しかしながら、現在の研究状況の下では、これら遺伝子は容易に入手できない。従って、C M C * 表現型を与える Clostridium thermocellum セルロース (c e l A) 用の c e l A 遺伝子が好ましく使用されるが、他のマーカー遺伝子、特に E. coli の lac Z を使用することも可能である。

マーカー遺伝子が c e l A である場合、B a t X I のような適切な制限部位へのコード配列の挿入後、C M C * 特性のために形質転換バクテリアが選択される。

本発明の方法においては、特にマーカー遺伝子とコード配列との間に制限部位を配置することにより、構造の点換後にマーカー遺伝子を容易に除去できることが好ましい。

コード配列は自然のもの、合成のもの、またはこれらの中間のものでもよい。

本発明の発現および分離系は、もちろん工業的に重要な生成物の生成を保証するように特別に設計されている。

従って、コード配列は、工業的に重要なペプチド、ポリペプチドまたは蛋白質を特別にコードする。しかしながら、このコード配列は、工業的に重要な蛋白質を直接的にコードするのではなくて、工業的に重要なアミノ酸、ポリペプチドまたは蛋白質の成熟および／または生成を必要とする蛋白質をコードする配列であってもよい。

本発明の方法はアミノ酸配列、特に反復配列の発現のために特別に設計されており、従って、これらは主に合成配列である。

これら固々の生成物をコードするこの第二 DNA は、また分離生成物の成熟を保証するように設計された特定の要素を含んでいてもよい。

合成配列の場合、コード配列を選択することにより次の構成が得られる。

- アミノ酸配列：

- (A A A ... A A A)_n 以上の n 個の反復単位を有する反復アミノ酸配列：

- C O O H - 宿主位置 A_i に正または負に带電したアミノ酸を含む反復配列。このアミノ酸は遺伝子発現を改善するが、次の事項を有利に達成し

(I) 喜しいイオン性によりポリペプチドを離離すること；

(II) 特定の蛋白質によりポリペプチドを (A A A ... A A A)_n 単位に切断すること；

特表平6-502548 (5)

(III) 必要ならば、特定のカルボキシペプチダーゼにより末端アミノ酸 α_1 を除去すること。
一所定の利点を有するアミノ酸を NH_2 - または COOH - 末端部分 α_1 または α_1 に含有する反復配列。実施例において、発現配列は構造 ($\alpha_1 - gln$)₁₀ および ($\alpha_1 - gln - lys$)₁₀ のポリペプチドをコードする。 $\alpha_1 - gln$ または $\alpha_1 - gln - lys$ 配列はその後の課素処理により放出することができる。課素処理または化学的処理により放出できる Ala-Gln-Tyr または Ala-Gln-Met のようなこの種の他のポリマーを生成することもできる。

コード配列のコドンの選択は、コリネバクテリア中の発現に影響を及ぼし、約 50% - 60% の GC 含有量を有する配列を生成することが好ましい。

この例の場合、(AQ)₂₀をコードする配列は、GCX-CAG であり、ここで X は A, T, C または G であり、実際これらコドンはアラニンに対して好ましくないが、一方 CAG コドンはグルタミンに対して明らかに好ましい。この場合、GC の含有量は約 75% であり、これは限定期である。従って、含有量を 55% まで減少させるために、第 3 番目は A および T が豊富な 3 つのアミノ酸を含むポリマーの使用が考えられる。

Tyr, Lys および Met はこれらコドンの最初の

2 つの塩基中に 2 つの A または T を有しており、従って、GC の含有量は 75% から約 60% に減少し、コリネバクテリア中に見られる GC の含有量に接近する。さらに、もちろん、グルタミン (Q) の COOH - 末端位置における工業的に重要なこれら 2 つのアミノ酸は、開発され且つ実在するものである。

本発明は、また上記免現および分詮系を含むコリネバクテリア菌株に関しており、特に、この場合、上記菌株は Brevibacterium, 特に Brevibacterium lactofermentum 菌株である。

最後に、本発明は、アミノ酸、ポリペプチドまたは蛋白質を生成する方法であって、上記コリネバクテリアの形質転換株を培養液中で培養することからなり、第二 DNA 配列が上記アミノ酸、ポリペプチドおよび / または蛋白質をコードし、且つ上記生成物が培養後に培養液から任意に分離されることを特徴とする方法に関する。実際、この方法により、有用な生成物が分離された。従って、この生成物は、この生成物自身を公知の方法で単離することのできる培養液中に存在する。これら公知の方法は例えばクロマトグラフィーまたは選択性沈殿のような分離方法であり、この選択性沈殿は、生成する分子の特性に適応したものであることが明らかに必要である。

また、バクテリア濃縮液を分離し、つぎに、例えば表面活性剤を用いて、この濃縮液から PS 1 または PS 2

（硫酸）および g_d_h_A を有する系用の糖分（グルコース / フルクトース）によって調整される。

また、本発明は、SP 1 または PS 2 配列のすべてまたは一部を含む蛋白質、特にこれら蛋白質の 1 つ以上の抗原部位を含む蛋白質に関する。上記蛋白質は、また代表的な要素として、特に診断セットとして、対応する抗体と使用することができる。

また、本発明は、所定の蛋白質が下記固定作用を行なう SP 1 または PS 2 部分の壁に固定され、また SP 1 または PS 2 の抗原エピトープが壁上に露呈されているコリネバクテリア菌株に関する。

下記の実施例は本発明の他の特徴および利点を示すことを意図しており、これら実施例はけっして本発明を制限するものではない。

第 1 図はプラスミド pCGL612 の図であり、このプラスミドは、蛋白質 PS 1 を合成する全 c-p1 遺伝子を含む C. melassecola ATCC17965 の 2, 6-トキフランメントを含む pUN121 (Hilssen, B., Uhien, K., Josephson, S., Catzenbeck, S., および Phillipson, L. (1988), An improved positive selection plasmid vector constructed by oligonucleotide mediated mutagenesis (オリゴヌクレオチド仲介突然変異誘発により生成された改良した正の選択プラスミドベクター)、Nucleic Acids Res 11: 1019-1029) から譲

と融合または別の方法で結合した所定の蛋白質を分離することも可能である。実際、PS 1 および PS 2 は壁蛋白質であり、このシステムにより分離した蛋白質の一部は、壁に固定したままであり、このことは、これら蛋白質の分離を容易にする。なぜならばバクテリアは特定の洗浄剤では溶離しないからである。

プラスミドによるコリネバクテリアの形質転換は、エレクトロポレーション (Bonamy C., Guyonvarch A., Reyes, O., David F. および Labios G. PEWS Microbiology Letters, 68, 263-270, (1990)) または他の好適な方法により行なうことが好ましい。

アミノ酸、ペプチドおよび / または蛋白質の生成を可能にする免現条件は、得られた生成物の種類並びに使用した特定の菌株に明らかに依存し、当業者の知識に従って各菌株に対して明確に決定しなければならない要素がある。

また、本発明は、csp1, csp2 および gdhA、これら 3 つの遺伝子のすべてまたは一部の発現用シグナルのすべてまたは一部を含む免現系、並びにこの種の系を発現する菌株、特にコリネバクテリアの菌株に関する。

上記構成物を使用する方法において、所定のアミノ酸、ポリペプチドまたは蛋白質の発現 / 分離は、温度、培養液および / または SP 1 および PS 2 用の糖分の性質、および pH (特に NH_4^+) の濃度、代謝産物 (グルタミ

導されたものである。

第2図は、*Corynebacterium melassecola* ATCC 17965と呼ばれている*Corynebacterium glutamicum*のcsp1遺伝子のスクレオチド配列および対応するアミノ酸配列を示す。スクレオチドの番号は、図面の右側に記載されている。反復スクレオチド配列は四角で囲ってある。予想されるSD配列には下線が付してある。転写ターミネーターにおそらくは対応する24-bpバリンドロームは、向かい合う矢印で示されている。この配列は、アクセス番号X66078の下でEMBLスクレオチド配列データバンクで見ることができる。

第3図はcsp1を保有する*C. melassecola* ATCC 17965の配列DNA領域の制限地図である。

第4図は、*C. glutamicum*の蛋白質PS1および*Mycobacterium*の抗原85複合体の蛋白質の配列の並列状態を示している。85B M. K. は*M. kansaii* (XIPSG18235) の抗原85-Bを表わす。85B M. b. は*M. bovis* (XIPSC83L79) の抗原85-Bを表わす。85B M. l. は*M. leprae* (EMBLX80934) の抗原85-Bを表わす。85C M. t. は*M. tuberculosis* (ENBLX57229) の抗原85-Cを表わす。85A M. b. は*M. bovis* (XIPSA28544) の抗原85-Aを表わす。85A M. t. は*M. tuberculosis* (XIPSI80082) の抗原85-Aを表わす。配列は「Genetics Computer

Group J (米国、ケイスコンシシ大学) のFastAプログラムを用いて整列した。長基の数は各行の始めに各蛋白質に対して与えられている。異なる各蛋白質の間に見られる同様のアミノ酸残基は四角で囲ってある。同様であると考えられる残基は次の通りである。酸またはアミド (D, E, N, Q) ; 塩基 (H, K, R,) ; 極性 (P, A, C, S, T) ; 無極性 (I, L, M, V) および芳香族 (F, W, Y) 。7つの蛋白質の間の同じアミノ酸残基は、関係した残基の上の星印によって表示されている。

注意： 各抗原に対して、アクセス番号は上記データバンクの名前と一緒にカッコ内に表示されている。

第5図はcsp1遺伝子の切断を示している。切断したcsp1遺伝子の染色体への組込み状態が示されている。pCGL613'は*C. glutamicum* 中で複製不可能であり、これはaphA3遺伝子 (I₀) によって切断されたcsp1 (黒色領域) を含んでいる。wtは*B. lactofermentus* 15野生型を表わし、Δcsp1は切断csp1を含む組み込み体である。

第6図はプラスミドpCGL616の構造を示している。プラスミドpCGL616は、*C. glutamicum*のcsp1遺伝子を有するプラスミドpCGL125に対応する。

第7図は切形蛋白質PS1を合成することのできるブ

ラスミドを示している。

ここでは、pCGL616から誘導したベクターの図、期待された蛋白質の大きさの明確、および抗PS1ポリクロナール抗体を使用しウエスタンプロットによる検出(+)または非検出(-)を示している。

第8図はプラスミドpCGL1030の構造を示している。図の領域Aはcsp1およびこれに隣くDNA領域を含んでおり、このDNA領域はPS1のシグナル配列およびその成熟配列の最初の30個のアミノ酸に対応している。

第9図はプラスミド1031の構造を示している。領域Aは第8図に説明されている。p+1およびECAの間の結合領域は連続配列されており、この配列の詳細が示されている。

第10図はプラスミド1032の構造を示している。領域Aは第8図に説明されている。PS1、(AQK)10およびECAの間の結合領域は連続配列されており、この配列の詳細が示されている。

第11図はプラスミド1033の構造を示している。領域Aは第8図に説明されている。PS1、(AQ)19およびECAの間の結合領域は連続配列されており、この配列の詳細が示されている。

第12図は、*Corynebacterium melassecola* ATCC 17965と呼ばれている*Corynebacterium glutamicum*

のcsp2遺伝子のスクレオチド配列および対応するアミノ酸配列を示す。スクレオチドの番号は、図面の右側に記載されている。予想されるSD配列には下線が付してある。転写ターミネーターにおそらくは対応する22-bpバリンドロームは、向かい合う矢印で示されている。

第13図はcsp2を保有する*C. melassecola* ATCC 17965の配列DNA領域の制限地図である。

第14図は*C. glutamicum*におけるcsp2遺伝子の切断を示している。切断された遺伝子の染色体組込み状態が示されている。*C. glutamicum* 中で複製不可能であるプラスミドpCGL830は、aphIIIにより切断されたcsp2遺伝子を有している。aphIIIおよびcsp2遺伝子の転写方向は、プラスミドpCGL830上の矢印によって表わされている。wtは*B. lactofermentus* 15菌株を表わし、csp2:aphIIIは切断csp2遺伝子を有する組み込み体である。

第15図は温度の関数としてPS1のトランスロケーションを示している。34°Cの指数増殖期 (OD650-1) における培養液の1.0mlを1分間³⁵Sメチオニン (37TBq/mmol, 1.6nM最終濃度) で標識した。パルスの終了時に、クロラムフェニコール (100μg/ml) および次に³²Sメチオニン (最終濃度0.5μM) を加えた。アリコートの1mlを除去し、所定の温度まで

特表平6-502548 (7)

他の可能な遺伝子 (DGF1、DGF2) との融合を可能にする celIA 遺伝子の 5'末端における合成配列

■ : ポリペプチドAQに等価なBstXI部位に導入したヌクレオチド (DGF5、DGF6)

□ : EGAのシグナル配列の一部に等価なDNAの配列

▨ : EGAのコード配列に等価なDNAの配列

△ : 転写ターミネーター

P : EGAのシグナル配列に属する最初のアミノ酸

第22図はpCGL125の構造を示す。

実施例1. 培養液上澄みおよびCorynebacterium glutamicumの壁におけるPS1およびPS2の同定

現在、*Corynebacterium glutamicum*菌株 (Jones, D., およびCollins, M.D. (1988), Irregular nonsporing Gram-Positive rods. In *Bergey's Manual of Systematic Bacteriology*, WilliamsおよびWilkins(eds.), Baltimore, 2巻, 1281-1434頁) と再定義されている *Corynebacterium celassaecola* ATCC17965菌株の培養液上澄みの変性条件 (SDS-PAGE) (Laemmli, U.K. (1970), Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*, 227: 680-685) の下におけるボリアクリルアミドゲル分析によれば、分子量がそれぞれ

速やかに冷却した。この温度で30分間培養を続け、PS1の分泌型留分を抽出した。次に、この抽出物を SDS-PAGEおよびオートラジオグラフにかけた (a)。バンドの強度は、デンシメトリーにより決定され (b、左側の軸) 且つ34℃で100を基準にして任意の単位で表わされている。トランスロケーションは蛋白質の相転移の開数である。

第16図はgdhA遺伝子の制限地図である。

第17図は*C. celassaecola*のgdhA遺伝子を含むNheI-BglIフラグメントの完全な配列を示す。

第18図はpCGL141およびpCGL142の構造、およびgdhA遺伝子のプロモータとlacZ遺伝子との間を融合するベクターを示す。

第19図は各構造において用いられるオリゴヌクレオチドを示す。

第20図はpPROK(AQ)₂₀celIAの構造を示す。

第21図はptacとcelIAとの間に合成遺伝子が配置されている構造を詳細に示しており、a)はptacの指令下にcelIAを配置した構造であり、b)はポリペプチドAQの導入後の予想できる構造である。

ptac : tacプロモータ

RBS : リポソーム結合部位

▨ : ポリペプチドAQに等価な配列の導入および

約67000および63000のPS1およびPS2の2つの主要蛋白質が明示される。PS1およびPS2の濃度は、バクテリアの成長曲線に従い、これらの定常期相において最大値に達する。これら蛋白質、特にPS2の大部分の分泌型留分は、バクテリアの壁中にも存在している。壁からPS1およびPS2を抽出するためには、バクテリアを殆ど溶解することのないバクテリアのSDS処理が使用される。従って、PS1およびPS2の最大濃度を得るために、2つの分泌型留分、培養液上澄みおよび細胞型留分を審査し、PS1およびPS2が高い割合で存在する最終標品を得ることができる。ポリクロナール抗体はPS1およびPS2に対して生成され、2つの蛋白質間に免疫交叉反応は存在せず、このことは、これら蛋白質が異なることを効果的に示している。

PS1およびPS2との強い免疫交叉反応性を有する蛋白質は、*Brevibacterium lactofermentum* 15 (Bonassasi, S., Oreglia, J., Trautvetter, A., およびSicard, A.M. (1990), Isolation and characterization of a restriction and modification deficient mutant of *Brevibacterium lactofermentum*, PEWS Microbial Letter 72:143-146)、*Brevibacterium lactofermentum* ATCC21086および*Brevibacterium flavum* ATCC14067菌株のような*Corynebacterium celassaecola* ATCC17965に関するバクテリア菌

株の培養液上澄み中に見いだされた。PS1およびPS2は、インベルターゼ、ベクチナーゼ、ヌクレアーゼ、コラゲナーゼ、アミラーゼ、バクテリオシン、エンドグルカナーゼおよび広域スペクトルプロテアーゼ活性を含む幾つかの酵素活性についてテストされた。これらの酵素活性はPS1またはPS2に包含されていなかった。

実施例2. *Escherichia coli*におけるPS1のシグナルペプチドの作用についての証明 (第1図)

プラスミドpCGL612 (第1図) により保持されたcp1遺伝子E. coli TG1中で発現されると、抗PS1抗体とこの粗抗体の粗抽出物のウエスタンプロット分析により、*C. celassaecola*の培養液上澄み中に存在する蛋白質PS1と同じ分子量を有する主要蛋白質の存在が明らかになる。わずかに大きな分子量を有する少量の蛋白質もまた検出される。実際、主要な蛋白質バンドは、PS1 (シグナル配列を有せず) およびPS1 (シグナル配列を有する) の前駆体型に対する副次的蛋白質バンドに対応している。

実際、第1の実験において、表面ショック (Reppel, L.A., (1987), Selective release of enzymes from bacteria, Science 158: 1451-1455) により破壊菌株 E. coli TG1 (pCGL612) のペリプラズミック蛋白質 (分泌した酵素) の放出および抗-PS1抗体使用のウエスタンプロット法による放出蛋白質含有

量の検出は、主要蛋白質のみを示す。この菌株のイソクエンチ酸塩デヒドロゲナーゼ活性 (Shilo, I., よび Ujigawa, K. (1978), Enzymes of the glutamate and aspartate synthetic pathways in a glutamate-producing bacterium, *Brevibacterium fluvium*, J. Bolchem 84: 647-657) は、溶菌をモニターする方法で測定した。この実験において、この溶菌は 1% 未満であると見抜もられた。このことにより、結論として主要蛋白質バンドは PS 1 の熟成型に対応し、且つこの蛋白質は *E. coli* の細胞膜を横切って放出される。

第 2 の実験において、粗抽出液 E. coli TG 1

(pCGL612) の粗抽出物は、蛋白質の合成を抑制するためにクロラムフェニコールの添加の前後に抗-PS 1 抗体を使用のウエスタンプロット法により分析された。副次的バンドは蛋白質合成の抑制後次第に消える。クロラムフェニコール CCCP (ローコロフュニルヒドロゾンシアン化カルボニル) の添加 5 分前に、細胞膜を横切る移動力を消散させるプロトノフォア (Protoaphorophore) が加えられるならば、この副次的 PS 1 のバンドは消失しない。従って、副次的 PS 1 のバンドの消失は、プロテアーゼによる分解の結果ではない。蛋白質合成の抑制後のその漸進的消失およびペリプラズムからのその不在は、膜内に位置されているペプチダーゼシグナル配列および細胞膜を横切るそのransporter

ーションによるこの前駆体形態の成熟の仮説に一致している。また、この結果は、*E. coli* において、PS 1 の成熟が生体内のプロトン移動力に依存していることも示している。

実施例 3. PS 1 をコードする csp1 遺伝子のヌクレオチド配列

上游領域の csp1 と呼ばれている PS 1 - コード遺伝子を含む 2547 塩基対 フラグメントの配列決定を行った。ヌクレオチド配列は第 2 図に示されている。(配列番号: 1)。第 3 図はこの配列領域の制限地図を表している。

コンピュータ分析を行なうことにより、657 個のアミノ酸に対応する 1971 個の塩基対の読み取り枠が確認された。

翻訳 (GAGAAGGAAAAACTTCATG) および転写 (TACATA (-35) および TAAAGAT (-10)) を開始する推定シグナルが確認された。上記リボソーム - 結合部位から抽出した AGAACGGA 配列は、グラム陽性型バクテリア、*Staphylococcus aureus* および *Streptomyces lividans*

(5' - G A U C A C C U C C U U U C U O H - 3') の rRNA の 3' 端部に対して相補的である (アンダーライン) (McLaughlin, J.R., Murray, C.L., および Rabieovitz, J.C. (1981), Unique features of the

ribosome-binding site sequence of the Gram positive *Staphylococcus aureus* β -lactamase gene., J. Biol. Chem. 256: 11283-11291) (Bibb, M.J. および Cohen, S.N. (1982), Gene expression in *Streptomyces*: Construction and application of promoter-probe plasmid vectors in *Streptomyces lividans*, Mol Genet 187: 265-277)。翻訳開始コドンの前方の 5' における DNA 領域は、2 つのヌクレオチド配列 AAAAGTTATCCACAG および ATTGAAA AA を含んでおり、これら配列のそれぞれは、最初の場合 28-42 および 70-84 において、二度目の場合 100-108 および 171-179 において二度繰り返されている。これら 2 つの配列は csp1 遺伝子の転写の開始範囲内に含まれる。

分泌シグナルの場合、蛋白質の NH₂ 末端における配列は、グラム陽性型バクテリアのシグナル配列の特徴を表わしている (Watson, M.E.E. (1984), Compilation of published signal sequence, Nucleic Acids Res., 12: 5145-5184)。このシグナル配列は、NH₂ - 末端位置における過剰の正電荷 (最初の 18 個のアミノ酸中に正電荷を有する 7 個のアミノ酸)、それに続いて過剰の無電性アミノ酸 (次の 23 個のアミノ酸中に 18 個のアミノ酸) を有する配列、さらにそれに続くシグナル配列切断部位の 2 つの推定アミノ酸配列 (28-32 の位置に

おける pro - thr - ala - ilo - ala) (39-43 の位置における pro - met - ala - ser - ala) を含んでいる。シグナル配列切断部位のこれら推定アミノ酸配列の間で、後者の配列 - 39-43 の位置における pro - met - ala - ser - ala は、最も確立が高いように思える。実際、蛋白質 PS 1 は、2 つの異なる方法 (実施例 5 を参照) を用いて電気泳動の結果が同じになるまで、*Corynebacterium glutamicum* の培養液上澄みから精製され、得られた標品は、エドマン分解法によりアミノ - 末端配列を決定するために用いられた。5-neo の純粋な蛋白質を使用したが、シグナルは得られなかった。2 つの精製方法が使用されたので、蛋白質 PS 1 が生体内で切断され、この切断は使用した精製技術の結果ではないと思われる。提呈された第 2 の切断配列は、成熟配列の最初のアミノ酸としてグルタミン (44 の位置) を明示していると思われ、このグルタミンはエドマン技術によりビログルタミン酸に容易に交換されて、蛋白質のアミノ - 末端配列化を可能にする。

rh o - 依存型の推定ターミネーター部位は、3 つの停止コドン och - amb - opa から 55 個のヌクレオチドの遺伝子の 3' 領域中に見いだされる (Rosenberg, M. および Court, D. (1979), Regulatory sequences involved in the promotion and termination of RNA

特表平6-502548 (B)

実施例4. *Corynebacterium glutamicum*のPS1と
*Mycobacterium*の抗原85複合体の蛋白質との間の配列
相同意性(第4図)

蛋白質PS1のNH₂部分は、3つの分泌ミコバクテリア抗原85-A、85-Bおよび85-Cにかなり類似している(Closs, O., Harboe, M., Axelsen-Christensen, H.H., and Magnussen, H.(1980) The antigens of *Mycobacterium bovis*, strain BCG, studied by crossed immuno-electrophoresis: a reference system Scand J Immunol 12: 249-263.) (Viker, R.G., Harboe, M., Nagai, S., and Bøgeadsen, J. (1990) Quantitative and qualitative studies on the major extra-cellular antigen of *Mycobacterium tuberculosis* B37Rv and *Mycobacterium bovis* BCG. Am Rev Respir Dis 141: 830-838.)。異なるミコバクテリア種類の3つの対応する遺伝子がクローニングされ且つ配列された。即ち、*Mycobacterium bovis* BCG 1173P2および*Mycobacterium tuberculosis*の抗原85-A(Borremans, H., De Wit, L., Volckaert, G., Doss, J., De Bruyn, J., Buyken, K., Van Vooren, J.-P., Stelandre, H., Verhofstadt, R., and Content, J. (1989) Cloning, sequence determination, and expression of a 32-kilodalton-protein gene of *Mycobacterium tuberculosis*. Infect Immun 57:

transcription. Annu Rev Genet 13: 319-353)。このヘアピン構造のΔGは-3.5, 7-kcal/molに等しい(Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T.およびTurner, D.H. (1988) Improved free-energy parameters for predictions of RNA duplex stability. Proc Natl Acad Sci USA 83: 9373-9377)。

読み取り枠中に含まれる657アミノ酸に対応する計算上の分子量は70874。しかしながら、最も可能性の高いシグナル配列(アミノ酸の42と43との間の切断部位)の分子量は4411であり、このことは成熟蛋白質に対して66463の計算上の分子量が与えられ、この値は、変性ポリアクリラミドゲルに基づいて計算された67000の値にかなり近似している。

配列の特徴をまとめると、次の通りである：

| | |
|-------------|--------------------------|
| 239 ~ 244 | TACATA (シグナル - 35) |
| 269 ~ 274 | TAAGAT (シグナル - 10) |
| 405 ~ 414 | GAGAAGGAAA リボソーム結合部位) |
| 420 ~ 2390 | コード配列 |
| 420 ~ 548 | 分泌蛋白のペプチドシグナル |
| 2455 ~ 2506 | ヘアピン構造、rho-型のターミネーターシグナル |

3123-3130.) (De Wit,L., De la Cuvellerie, A., Doss, J., and Content, J. (1990) Nucleotide sequence of the 32 kDa-protein gene (antigen 85A) of *Mycobacterium bovis* BCG. Nucleic Acids Res 18: 3995.)、*Mycobacterium bovis* Tokyo、*Mycobacterium kansassii* および*Mycobacterium leprae*の抗原85-B(Matsu, K., Yamaguchi, R., Yamaasaki, A., Tasaka, H., and Yamada, T. (1988) Cloning and expression of the *Mycobacterium bovis* BCG gene for extracellular α antigen. J. Bacteriol 170: 3147-3154.) (Matsu, K., Yamaguchi, R., Yamaasaki, A., Tasaka, H., Terashita, K., and Yamada, T. (1990) Cloning and expression of the gene for the cross-reactive α antigen of *Mycobacterium kansassii*. Infect Immun 58: 550-558.) (De Mendonca Lima, L., Content, J., Van Heuverswyn, H., and Degrave, W. (1991) Nucleotide sequence of the gene coding for the 85-B antigen of *Mycobacterium leprae*. Nucleic Acids Res 19: 5789)、および*Mycobacterium tuberculosis*の抗原85-C(Content, J., De La Cuvellerie, A., De Wit, L., Vincent-Levy-Prebault, V., Doss, J., and De Bruyn, J. (1991) The genes coding for the antigen 85 complexes of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG

are members of a gene family: cloning, sequence determination, and genomic organization of the gene coding for antigen 85-C of *M. tuberculosis*. Infect Immun 59: 3205-3212.) *Corynebacterium glutamicum*の蛋白質SP1は、同じ残余α-1, 1の約33%を有する) および約330個のアミノ酸(+/-5)の長さにわたってこれら6個の蛋白質を有する同様の簇部(α-1, 1)の約52%。約330個のアミノ酸のこの長さは、ミコバクテリア抗原の場合、蛋白質の全長に対応する。ちょうどPS1のように、すべてのこれらミコバクテリア抗原は、グラム陽性型バクテリア(約42個アミノ酸、α-2, 4)中に見いだされる最も長いシグナル配列に匹敵する長さのシグナル配列を含んでいる。ちょうどPS1のように、*M. bovis*の蛋白質85-Bおよび*M. tuberculosis*の蛋白質85-Cは、大部分のシグナル配列より長い親水性NH₂領域(5個以上の正に荷電した残基)を有する。すべてのこれらのシグナル配列の他の重要な特性は、グルタミン酸である*M. tuberculosis*の抗原85-Cを除いて、酸残基、即ちアスパラギン酸の3または5の位置に存在する。酸性荷電残基の存在は、Eucaryoticシグナル配列のNH₂端部に共通なものであるが、原核シグナル配列のNH₂端部に対して全く例外的である。(Perlman, D., and Halvorson, H.O. (1983) A putative signal peptid

asrecognition site and sequence in eucaryotic and prokaryotic signal peptides. J Mol Biol 167: 391-409. (Watson, H.E.E. (1984) Compilation of published signal sequences. Nucleic Acids Res. 12: 5145-5164). この特徴の理由は知られていない。他の重要な観察点は、EMBL/MIPSデーターベンクに存在するPS1と他の蛋白質との間に見いだせない。

実験例5. N末端配列の決定に用いられるPS1及びPS2精製操作

操作1:

タンパク質PS1及びPS2をポリアクリルアミドゲルでの調製電気泳動及び電気溶出によりC.グルタミクムATCC17965の培養上液から精製した。

34°Cにおいて富士B培地200mlで培養された細菌を4°Cで15分間8000回転で遠心することにより定期増殖期に収集した。次いで培養上液のタンパク質を60%硫酸アンモニウムで沈降させ、4°Cで15分間13000回転で遠心することにより収集した。ペレットをpH6.8の10mMトリスHCl緩衝液4mlに溶解し、しかしる後溶液をこの同緩衝液中4°Cで24時間透析する。

硫酸アンモニウム沈降後に得られた透析タンパク質抽出液をフォーマット16×20×0.75cmの電気泳動ゲル上に沈着させる。電気泳動は4%濃縮用ゲル及び7.5%分離用ゲルを用いてLaemmli (1970)により記載

る。更に、分離用ゲルを使用する前にブレランする。これらすべての予防処置はタンパク質のN末端を変えて、ひいてはこの末端をプロックする可能性があるラジカルの形成をできるだけ避ける目的で行われる。電気泳動終了後、タンパク質をPVDF膜上に転写する。この操作はpH8.0の50mMトリス、50mMホウ酸緩衝液中50V、60分間で行う。次いで膜を位置決定を可能にするアミドブラックで染色し、タンパク質PS1及びPS2に相当するバンドを切出す。次いでそのタンパク質バンドを脱染して、それらをN末端配列決定に用いた [Laemmli, U.K. 1970. Cleavage of structure proteins during assembly of the head of bacteriophage T4(バクテリオファージT4の頭部のアセンブリーにおける構造タンパク質の開裂). Nature, 227:880-885; Lee, C., Levin, A., Branton, D. 1987. Copper staining: a five minute protein stain for sodium dodecyl sulfate polyacrylamide gels(銅染色: ドデシル硫酸ナトリウムポリアクリルアミドゲル用の5分間タンパク質染色). Anal Biochem, 166:308-312].

実験例6. PS1と一緒に呼ばれるものはやPS1を含むしないコリネバクテリウム・グルタミクム株の産生(図5)
プレビバクテリウム・ラクトファーメンタム(Brevibacterium lactofermentum)15と称されるC.グルタミクム株(C. glutamicus)は大腸菌K12の改変

特表平6-502548(10)

された操作に従い行う。泳動は40mlで15時間にわたり行う。次いでゲルをLees (1987)により記載された操作に従い塩化銅で染色する。タンパク質PS1及びPS2に相当するタンパク質バンドを切出し、しかしる後完全に脱染する。次いでタンパク質を48ml、4°Cで5時間かけてゲルから電気溶出し、しかしる後pH6.8の10mMトリスHCl緩衝液で数回透析してから、いくつかのアリコート分面に分け、-20°Cで凍結する。精製収率は90%以上の程度で25%程度である。

操作2:

タンパク質PS1及びPS2を限外通過、電気泳動及びPVDF膜上での転写によりC.グルタミクムATCC17965の培養上液から精製する。

34°Cにおいて富士B培地で培養された細菌を4°Cで15分間8000回転で遠心することにより定期増殖期に収集する。上澄4mlをpH7.0の50mMリン酸緩衝液で50倍希釈し、しかしる後カットオフが30kDである限外通過膜で遠心する。このステップによれば80μlタンパク質抽出液を得ることが可能であり、しかしる後これを4%濃縮用ゲル及び7.5%分離用ゲルからなる電気泳動ゲルに沈着させる。電気泳動は下記修正を加えてLaemmliにより記載された操作に従い行う。ゲル及び泳動用緩衝液を調製するために用いられるすべての溶液を脱氣させ、これらに0.1Mチオグリコレートを含有させ

DNAに対して許容的であり [Bonassie-S., Oreglia, J., Trautvetter, A. 及び Sicard, A.M. (1990). Isolation and characterization of a restriction and modification deficient mutant of *Brevibacterium lactofermentum* (プレビバクテリウム・ラクトファーメンタム)の制限及び修正欠失変異体の単離及び特徴化). PEWS Microbial Letters, 72:143-148]、一方C.メラッセコラ(C. melassecola) ATCC17965と称されるC.グルタミクム株は大腸菌のDNAに関して非常に制限的な株である [Reyes, O., Guyonvach, A., Bonamy, C., Saiti, V., David, P. 及び Lebida, G. (1991). "Integron"-bearing vectors: a method suitable for stable chromosomal integration in highly restrictive Corynebacteria(「インテグロン」-保持ベクター: 高制限コリネバクテリアにおける安定的染色体組込みに適した方法). Gene, 107: 1-68]。この理由から、B.ラクトファーメンタム15株をcsp1遺伝子の切断を実施するために選択した。csp1遺伝子の物理的地図はC.メラッセコラATCC17965及びB.ラクトファーメンタム15の場合で同一であることが確かめられた。

カナマイシン耐性(Km^r)を付与するストレプトコッカス・ファエカリス(Streptococcus faecalis)の△pHA3遺伝子を含むプラスマドpAT21の1.5kbCLal断片(Traieu-Cuot, P. 及び Courvalin, P. (1983).

Nucleotide sequence of the *Streptococcus faecalis* plasmid gene encoding the 3'-5'-aminoglycoside phosphotransferase type III (3'-5'-アミノグリコシドホスホトランスクフェラーゼタイプIIIについてコードするストレプトコッカス・ファエカリスプラスミド遺伝子のヌクレオチド配列). *Gene*, 21:331-341) をプラスミド pCGL612 中に存在する csp1 遺伝子の独特な KpnI 部位 (A sp718) に挿入して、プラスミド pCGL613'を得た。プラスミド pCGL613'を保有する組換え大腸菌株が実際に PS1-表現型であることは抗PS1ポリクローナル抗体を用いたウェスタンプロットティングにより示された。このプラスミドは C, グルタミクムではなく大腸菌で典型的に複製する。それを電気的形質転換により B, ラクトファーメンタム 15 と称される C, グルタミクムの株に導入し (Bonamy, C., Guyonvarch, A., Reyes, O., David, P., 及び Leblon, G. (1990). Interspecies electro-transformation in *Corynebacteria* (コリネバクテリアにおける種間電気形質転換). *FEMS Microbial Letters*, 88:285-270) 、 Km' 形質転換株を選択した。Km' 形質転換株において、プラスミド pCGL613' は宿主ゲノムの csp1 領域との相同的組換えにより C, グルタミクムの染色体中に組込まれるとと思われる。形質転換株の 22.5%において二重組換え現象が生じて、形質転換株プラスミドの csp1 :

p h A 3 相立体による野生型 csp1 遺伝子の置換を起こし、Km' - Tet' 形質転換型をえた (図 5)。BglII 又は BamHI 及び EcoRI のいずれかで切断された野生型株と Km' - Tet' 形質転換株の 1 つとの全染色体 DNA を pCGL613' プローブとのサンプルプローティングにより分析した (Sambrook, J., Fritsch, E. F., 及び Maniatis, T. (1989). *Molecular cloning: a Laboratory manual* (分子クローニング: 実験マニュアル), second edition. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Publications)。csp1 遺伝子は野生型株で約 7.5 kb 断片に含まれ、一方組込体 pCGL613' は csp1 遺伝子中に挿入された 1.5 kb spha3 遺伝子に相当する約 9 kb 断片を含む。BamHI - EcoRI 切断から図 5 で示された組込体の構造を確認する。

Km' - Tet' 組込体を抗PS1ポリクローナル抗体を用いて PS1 の発生に関するウェスタンプロットティングでも分析した。この株の培養上液又は粗製抽出物のいずれにもタンパク質 PS1 はない。これは Lgt11 に組込んでクローニングされた csp1 遺伝子が C, グルタミクムで実際に PS1 をコードする独特な遺伝子に相当することを確認させる。

この PS1-C, グルタミクム株は完全に生存可能であり、その増殖速度は影響をうけないようである。この

結果は痕跡的にみて生存力に影響を与えることなく C, グルタミクム株中への相同的又は異種 DNA の組込み用のターゲットとして csp1 遺伝子領域を使用できることを示す。

実施例 7. マルチコピーで csp1 遺伝子の C, グルタミクム中における発現、その合成及びその分泌に必要な PS1 の重要な領域の分析

この一連の実験のために、プラスミド pCGL616 を構築した。それは全 csp1 遺伝子を含んでおり、プラスミド pCGL1'25 から構築され、これは C, グルタミクムで複製でき、ストレプトコッカス・ファエカリスの spha3 遺伝子を含むクローニングカセットを備えたプラスミド pBL1 (Santamaría, R., Gil, J. A., Micas, J. M., 及び J. P. Martín (1984). Characterization of an endogenous plasmid and development of cloning vectors and a transformation system in *Brevibacterium lactofermentans* (ブレビバクテリウム・ラクトファーメンタムにおける内在プラスミドの特徴化とクローニングベクター及び形質転換系の開発). *J. Gen. Microbiol.*, 130:2237-2248)、及び大腸菌で複製でき、csp1 遺伝子を含むプラスミド pCSP1G に相当する。この組立てから得られるプラスミド pCGL616 (図 6) は C, グルタミクムで複製できる。

C, グルタミクム PS1 株における PS1 合成の回復

PS1 合成のいわゆる B, ラクトファーメンタム 15 PS1 株における回復はプラスミド pCGL616 の後者への導入後に観察される。プラスミド pCGL616 を保有する C, ラクトファーメンタム 15 PS1 株において、多量に分泌された PS1 は野生型 B, ラクトファーメンタム 15 株 (天然 PS1+) との比較により検出される。これは csp1 遺伝子のコピー数を増加させることにより C, グルタミクム株で分泌される PS1 の濃度を増加させることを示す。

この結果はいわゆる C, メラッセコラ ATCC 17965 株でも証明される。

端部切取り PS1 タンパク質の合成を可能にする pCGL616 に由来するプラスミドの構成 (図 7)

この実験は天然タンパク質に関する 67000 (Nv) の代わりに約 23000 (Nv) に相当する分子量の端部切取り PS1 タンパク質が C, グルタミクムでなお分離されることを示す。

7種の欠失をプラスミド pCGL616 から発見して csp1 遺伝子領域で行い、7種の異なるプラスミドを得た。これらすべての欠失は PS1 のシグナル配列に相当する DNA 領域と csp1 遺伝子の転写ターミネーターを保存している。すべての場合において、端部切取り PS1 タンパク質の合成及び分泌を抗PS1ポリクローナル抗体を用いたウェスタンプロットティングにより分析

ため、成熟 P.S.1 配列の 30 番目のアミノ酸の直後におりた。結果的に、このアラニンは分泌に必要な P.S.1 の培養液を離れており、したがって発現及び分泌双方の手段に対応している。

コリネバクテリウム・グルタミクムにおけるクロストリジウム・サーモセルムの c.e.1.A 遺伝子の発現及び対応タンパク質の分泌 (図 9)

エンドグルカナーゼ A 又は E.G.A と称されるエンドグルカナーゼについてコードする C. サーモセルムの c.e.1.A 遺伝子 (Corset.P., Hillet.J., Beguin.P. 及び J.P. Aubert (1983). Characterization of two cel (cellulose degradation) genes of Clostridium thermocellum coding for endoglucanases (エンドグルカナーゼについてコードするクロストリジウム・サーモセルムの 2 つの cel (セルロース分解) 遺伝子の特徴化). Bio/Technology, 1:589-594] を S.malI 部位でベクター p.CGL1.030 に組込んでクローニングし、プラスミド p.CGL1.031 を得た (図 9)。この c.e.1.A 遺伝子は、キメラ構造の目的でタンパク質 E.G.A の翻訳開始部位の非常に近くに Bst XI 制限部位を人工的に導入したプラスミド p.CGL1.008 に由来する (図 10, 11 参照)。タンパク質 E.G.A の合成は C.M.C と称されるエンドグルカナーゼ基質カルボキシメチルセルロースを用いたディッシュにおける酵素活性に関する染色試験によ

した。これらの結果は c.s.p.1 遺伝子のはほとんどを欠失しても、端部切取りタンパク質 P.S.1 の合成及び分泌をおこ可能であることを示している (図 7)。この結果を許容する最大欠失は c.s.p.1 遺伝子の約 1.3 kb の N.c.o.I - Bsp.E.I (Bsp.M.II) 断片の欠失に相当するが (p.CGL1.041)、これは分泌成熟形として約 2.9 kD 及び約 2.4 kD の P.S.1 に関する前駆タンパク質サイズを与える。約 2.3 kD のタンパク質はこのプラスミド p.CGL1.041 を保有するいわゆる B. ラクトファーメンタム 1.5 P.S.1-株の培養上澄で抗 P.S.1 抗体を用いてウェスタンプローティングにより実際には検出される。

実施例 8. c.s.p.1 系に基づき p.CGL1.030 と組まれる C. グルタミクムにおける発現及び分泌ベクターの構造 (図 8, 9, 10, 11)

プラスミド p.CGL1.030 の構造 (図 8)

C. グルタミクムで複製できるこのプラスミド (C. グルタミクムのプラスミド p.B.L.1 を含む) は、C. グルタミクムの c.s.p.1 遺伝子のプロモーターとシグナル配列に成熟 P.S.1 配列の最初の 30 アミノ酸を加えたものに相当するこの遺伝子の DNA 領域を保持している。マルチクローニング部位 (図 8 におけるポリリンカー 2) は、C. グルタミクムで発現されることを要するあらゆる非相同的遺伝子と同期して容易にクローニングさせる

り容易に検出できる [Corset.P., Hillet.J., Beguin.P. 及び J.P. Aubert (1983). Characterization of two cel (cellulose degradation) genes of Clostridium thermocellum coding for endoglucanases. Bio/Technology, 1:589-594]。この C.M.C 試験を C. グルタミクムにおける C. サーモセルムのタンパク質 E.G.A の合成を確認するために使用する。富培地 (LB + ルリアプロス又は B.H.I. - 脳心臓インキュベーション) 中全細胞又は培養上澄で実施されるディッシュでの活性に関する C.M.C 試験では、双方の場合においてプラスミド p.CGL1.031 を保有するプレビバクテリウム・ラクトファーメンタム 1.5 とみられる C. グルタミクム株のエンドグルカナーゼ活性を表す。より高い活性は LB + フルクトース又は + グルコース培地でみられ、これは c.s.p.1 プロモーターの調節下での c.e.1.A の発現へのこれら 2 種の糖の刺激効果を示している。これはザイモグラム (Beguin.P. (1983). Detection of cellulase activity in polyacryl amide gels using Congo red stained agar replicas (コンゴレッド染色寒天レプリカを用いたポリアクリルアミドゲルでのセルラーゼ活性の検出). Anal. Biochem., 131:333-338) 及び抗 E.G.A ホリクローナル抗体と共に培養上澄について行われるウェスタンプローティングで確認される。

合成ポリペプチド (A.Q.K) 1.0 の発現及び分泌に関する c.s.p.1 系の使用 (図 10)

10 回反復されるポリペプチドアラニン・グルタミン・リジンに対応する合成遺伝子を化学的に合成し、プラスミド p.CGL1.008 の Bst XI 部位に組込んでクローニングし、プラスミド p.CGL1.017 を得た。プラスミド p.CGL1.017 の Eco RI 断片を、シグナル配列の c.s.p.1 プロモーター及び P.S.1 の最初の 30 アミノ酸の下流 (及びリポーター遺伝子 c.e.1.A の上流) に位置するプラスミド p.CGL1.030 の S.malI 部位に組んでクローニングし、プラスミド p.CGL1.032 を得た (図 10)。キメラタンパク質 P.S.1- (AQK) 1.0 - E.G.A の検出をディッシュ中での C.M.C 試験、ザイモグラム又はウェスタンプローティングにより前記されたように行う。

合成ポリペプチド (A.Q.) 1.0 の発現及び分泌に関する c.s.p.1 系の使用 (図 11)

20 回反復されるポリペプチドアラニン・グルタミン・リジンに対応する合成遺伝子を化学的に合成し、プラスミド p.CGL1.008 の Bst XI 部位に組んでクローニングし、プラスミド p.CGL1.002 を得た。プラスミド p.CGL1.002 の Eco RI 断片を、シグナル配列の c.s.p.1 プロモーター及び P.S.1 の最初の 30 アミノ酸の下流 (及びリポーター遺伝子 c.e.1.A の上流) に位置するプラスミド p.CGL1.030 の S.malI 部位に組込

んでクローニングし、プラスミド p CGL 1033 を得た (図 11)。キメラタンパク質 PS1 - (AQ) 19 - EGA の推出をディッシュ中での CMC 試験、ダイモグラム又はウェスタンプロットtingにより前記されたように行う。プラスミド p CGL 1033 中 B、ラクトファーメンタムにおけるその配列はコード配列 AQ の喪失を示した (B、ラクトファーメンタムでクローニング中 AQ₂₀ から AQ₁₉ への置換)。

この一連の実験では csp1 遺伝子のプロモーターがクロストリジウム・サーモセルムの異種 celA 遺伝子とキメラ構造 (AQK) 10 - celA 及び (AQ) 19 - celA を C、グルタミクム中で発現させることを示す。更に、これらの実験では PS1 の諸要素、この場合ではそのシグナル配列とその後の異種遺伝子の上流に位置する各成熟配列の最初の 30 アミノ酸、が対応産物を分離させることを示す。培地の効果とこの培地への難、この場合ではグルコース又はフルクトース、の添加又は非添加は対応産物の生産に影響を与える。特に、csp1 プロモーターの調節下 C、グルタミクム中において、EGA 又はキメラタンパク質 (AQK) 10 - EGA もしくは (AQ) 19 - EGA の產生は BH1 培地よりも LB 培地で高く、それは LB 培地においてグルコース又はフルクトースで高度に刺激される。C、グルタミクムの csp1 プロモーターは C、サーモセルムの天然 celA

IA プロモーターよりも強いことがわかる：実際には、天然 celIA プロモーターを含むいわゆるプラスミド p CGL 602 保有 B、ラクトファーメンタム 15 株は、celIA が C、グルタミクムの csp1 プロモーターの調節下にある、プラスミド p CGL 1031 を保有するこの回体よりも実質上小さなエンドグルカナーゼ活性を有する。

p CGL 1032 又は p CGL 1033 を含む異なる株の培養上液について行われたウェスタンプロット実験では、いくつかのタンパク質バンドが抗 EGA ポリクローナル抗体と反応することを示す。これらの異なるバンドはエンドグルカナーゼ EGA に特異的であり (コントロールに不存在)、そのタンパク質及びキメラタンパク質の分解産物におそらく相当する。しかしながら、それより高い分子量のバンドも実際に (AQK) 10 - EGA (p CGL 1032) 及び (AQ) 19 - EGA (p CGL 1033) で密接して (Nv(AQ) 19 - EGA > Nv(AQK) 10 - EGA) 離離される。

実験例 9. コリネバクテリウム・グルタミクムのタンパク質 PS2 についてコードする csp2 遺伝子のヌクレオチド配列 (図 12, 13)

csp2 と称される PS2 をコードする遺伝子を含む 2702 基基對断片とその上流領域の配列決定を行った。ヌクレオチド配列は図 12 に示される (配列番号版 2)。

図 13 はこの配列決定された領域の制限地図を示す。

コンピューター解析を用いて、1532 基基對断片が 510 アミノ酸に対応することを確認した。

シャイシ・ダルガルノタイプ配列 AAGGAG を翻訳開始コドンのすぐ上流で確認した (-12 ~ -17)。

そのタンパク質の NH₂ 末端においてグラム陽性菌の非常にありふれたシグナル配列が 30 アミノ酸で存在する。シグナル配列開始部位の推定アミノ酸配列 His pro ala phe ala が発見された。コリネバクテリウム・グルタミクムの培養上液から精製されたエドマン分解技術によるタンパク質のアミノ末端配列の決定では、5 mol の精製タンパク質を用いたけれども、シグナルが得られなかつた。2つの精製操作を用いたため、タンパク質 PS2 は PS1 と全く同様にインピゴでブロックされるが、そのブロッキングは用いられた精製技術の結果でないようである。30 アミノ酸に関して提議されたシグナル配列は成熟配列の最初のアミノ酸としてグルタミン (31 位) を現すが、これはエドマン技術によるタンパク質のアミノ末端配列決定を不可能にするビログルタミン酸に容易に変換される。このタンパク質 PS2 はその非常に酸性の特性 (pI = 4.1)、そのシステイン残基欠如及びその非常に低いメチオニン残基含有率のようなタンパク質の特徴を有する (Sleytr, U.B. (1978). Regular arrays of macromolecules on bacterial cell walls:

structure, chemistry, assembly and function (細胞壁膜における高分子の規則的配列：構造、化学、アセンブリー及び機能). Int. Rev. Cytol. 53:1-84) [Sleytr, U.B. 及び P. Nezzer (1983). Crystalline surface layers on bacteria (細胞壁における結晶表面層). Ann. Rev. Microbiol. 37:311-339]。電子顕微鏡分析では PS2 が実際に細胞表面において組織化された六方晶構造で自ら配列できる型タンパク質であることが確認されている。

△ 非依存性タイプの推定ターミネーター部位は停止コドンから 76 ヌクレオチドでその遺伝子の 3' 領域においてみられる。

その配列の特徴は下記のとおりである：

562 ~ 567 : リボソーム結合部位

579 ~ 2108 : コード配列

579 ~ 668 : 分泌タンパク質のシグナル配列

2188 ~ 2233 : ヘアピン構造、△ 非依存性タイプの推定ターミネーターシグナル

(停止コドンから 76 ヌクレオチドに存在)

実験例 10: PS2 と称されるもしくはや PS2 を合成しないコリネバクテリウム・グルタミクム株の產生 (図 14)

csp2 遺伝子の選択は、 λ phage の挿入により不活化された csp2 遺伝子のコピーを保持する、コリネバクテリアについての非演習的なベクター、ベクター p

C G L 8 3 0 (図14) (プラスミド p C G L 8 1 1 により保持される c s p 2 の独特な N r u I 部位に a p h III 遺伝子を組込んだクローニング) により B. ラクトファーメンタム 1 5 と称される C. グルタミクムで行った。P S 2 シグナルはプラスミド p C G L 8 3 0 を保持した大腸菌 T G 1 株に由来する細胞抽出物についての抗 P S 2 ポリクローナル抗体での免疫学的検出により示されなかった。組込まれたクローニングは B. ラクトファーメンタム 1 5 株のエレクトロポレーション及び K m に関する選択により選択した。これら組込みのうち、選択遺伝子により野生型遺伝子の置換を起こす二重交換現象を示す T c t ⁺ クローニングを得た。

プローブ p C G L 8 1 1 を用いた K m ^r T c t ⁺ の Xba I 及び Sac I 切断染色体 DNA のサンプルロット分析は野生型株で得られる 2. 7 kb Xba I 及び 0. 7 kb Sac I の代わりに各々 4. 2 kb 及び 2. 2 kb の断片を示すが、これは a p h III 遺伝子の存在と連関したサイズ増加を示す。

異なる分画において抗 P S 2 ポリクローナル抗体でのウェスタンプローティングによる P S 2 の検出の次如は、B. ラクトファーメンタムにおける c s p 2 遺伝子の選択を確認させる。この P S 2 - 株は完全に生存可能であり、決してその増殖に影響をうけない。c s p 1 遺伝子を保持する C. グルタミクムの染色体の領域と同様

に、c s p 2 遺伝子を保持するこの DNA 領域も細菌の一時宿主に影響を与えることなく外來 DNA の組込み用ターゲットとして使用できる。

B. ラクトファーメンタム 1 5 P S 2 - 株における P S 2 + 表現型の回復

全 c s p 2 遺伝子とその上流の DNA 領域を含む 2. 3 kb Sac I - F i p I 断片をプラスミド p C G L 8 2 4 に組込んでサブクローニングし、B. ラクトファーメンタム 1 5 P S 2 - 株に再導入し、P S 2 + 表現型を回復させた。より多量の P S 2 はその遺伝子がマルチコピーで存在する場合に得られることに留意すべきである。これらの結果は C. グルタミクムの c s p 2 遺伝子を由来する分離産物の量がその遺伝子のコピー数に従い改変できることを示す。

クリオフラクチャー (cryofracture) により (前記技術により得られた) 株 P S 2 + 及び P S 2 - のサンプルの電子顕微鏡分析は、タンパク質 P S 2 が細胞表面において組織化された六方晶構造で自ら効率的に配列できる整タンパク質であることを非常に明確に示す。

実施例 11. P S 1 の分泌に関する温度の効果 (図 15)

指致増殖期 (34°C) における細胞を ³⁵S メチオニンで 1 分間かけて標識した。次いでクロラムフェニコール (100 μg/ml) 及び過剰量の冷メチオニン (³²S) を加えた (時間 0)。次いで細胞懸濁液の温度を室温まで

温度まで速やかに調整し、インキュベートを上記温度で 30 分間続ける。P S 1 の移動を SDS-PAGE、オートラジオグラフィーにより調べ、密度計測により定量する (図 15)。P S 1 の移動は明らかに温度に依存している。移動は 10°C 以下で起きず、それはこの温度を超えて最大約 30°C に達するまで急速に増加する。移動は脂質の相転移と相關している (図 15)。

実施例 12. コリキバクテリウム・メラッセコラ

A' T C C 1 7 9 6 5 に関する染色体 DNA ライブラーの構成及び g d h A 遺伝子のクローニング

C. メラッセコラの株 A' T C C 1 7 9 6 5 の染色体 DNA を Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (Eds) ((1987) Current protocols in Molecular Biology (分子生物学における現行プロトコール). John Wiley and Sons, New York) により記載された方法に従い得た。制限エンドヌクレアーゼ M b o I (ベーリンガー (Boehringer)) による制御的切断を Maniatis, T., Fritsch, E. F., Sambrook, J. ((1982). Molecular cloning: a laboratory manual (分子クローニング: 実験マニュアル). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York) により記載された操作に従いこの DNA 10 μg で行った。DNA 断片を Ausubel ら (1987) で記載されたようにスクロース勾配でそれらのサイズに従い分離した。サイズ 6

~ 1.5 kb の断片をライブラー構築用に選択した。

クローニングプラスミド p U N 1 2 1 [Nilsson, B., Uhlen, M., Josephson, S., Catzenberg, S., Phillipson, L. (1983). An improved positive selection plasmid vector constructed by oligonucleotide mediated mutagenesis (オリゴヌクレオチド媒介変異誘発により組立てられた改良陽性選択プラスミドベクター). Nucleic Acids Res. 11:8019-8030] を Dr. B. Bachmann から自由に入手できる大腸菌の株 GM 2929 から Birnboim, H. C., Doly, J. ((1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA (組換えプラスミド DNA をスクリーニングするための迅速なアルカリ抽出操作). Nucleic Acids Res. 7:1513-1521) の方法により得た。そのプラスミドを新規エンドヌクレアーゼ B c 1 1 (ベーリンガー) で直鎖化した。

ライブルーは B c 1 1 で直鎖化されたプラスミド p U N 1 2 1 1 μg 及び前記 6 ~ 1.5 kb DNA 断片 2 μg の Ausubel ら (1987) により記載された条件下における T 4 DNA リガーゼ (ベーリンガー) での結合により構成した。結合混合物は Dover, V. J., Miller, J. P., Raghda, S. C. V. ((1988). High efficiency transformation of E. coli by high voltage electroporation (高電圧電気穿孔法による大腸菌の高効率形質転換). Nucleic Acids Res. 16:8127-8145) により記載された操作に従いエレクト

ロボレーションより大腸菌株 DH5 中に導入した。組換えプラスミド保有大腸菌クローニングをテトラサイクリン 10 μg/ml 含有 LB 培地上で増殖できるか否かにより直接選択した。全テトラサイクリン耐性クローンのプラスミドを Birnboim 及び Doly (1979) の方法により得た。これらプラスミドの組合せは DNA ライブライリーに相当する。

グルタミン酸デヒドロゲナーゼ活性を欠いた大腸菌 CLR 207 rec A [Mattaj, J. V., McPherson, M. J., Yotova, J. C. (1982). Localization of a strongly conserved section of coding sequence in glutamate dehydrogenase genes (グルタミン酸デヒドロゲナーゼ遺伝子におけるコード配列の強く保存されたセクションの局在化). FEBS Letters, 147: 21-25] を C. メラッセコラ ATCC 17965 DNA ライブライリーで形質転換した。アンビシリソ 100 μg/ml 含有最少選択培地上で増殖できる大腸菌 CLR 207 rec A の形質転換クローンを選択した。このクローンは組換えプラスミド pCGL310 を保持する。Meers, J. L., Tempest, D. V., Broova, C. H. [(1970). Glutamine (amide): 2-oxoglutarate amide transferase oxidoreductase (NADP)-an enzyme involved in the synthesis of glutamate by some bacteria (グルタミン (アミド) : 2-オキソグルタル酸アミノトランスフェラーゼオキシドレダクターゼ (NADP) -一部細菌によるグルタミン酸の合成に関与する酵素). J. Gen. Mi-

特表平6-502548 (16)

crobiol. 64: 187-194] の方法に従い測定されるグルタミン酸デヒドロゲナーゼ活性は、プラスミド pCGL310 保有大腸菌 CLR 207 rec A で回復される。我々のサブクローニングによって、第一に完全 gdhA 遺伝子を保持する C. メラッセコラの DNA 断片を Eco RI 及び Xba I 制限部位により範囲限定される 3.8 kb DNA 断片にまで短縮することができた。この Eco RI - Xba I 断片の正確な制限地図は 図 16 で示される。次のサブクローニングによって更に正確に gdhA 遺伝子を 2.2 kb Nhe I - Bgl II 断片にまで範囲限定することができた。Southern, B. M. [(1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis (ゲル電気泳動により分離された DNA 断片の中における特異的配列の検出). J. Biol. Biol. 98: 503-517] の方法による DNA - DNA ハイブリッド形成は、クローン化 DNA 断片が実際に C. メラッセコラの株 ATCC 17965 に由来することを示した。

gdhA 遺伝子のヌクレオチド配列の決定

前記 Eco RI - Xba I DNA 断片のヌクレオチド配列の決定を実施するため、下記サブクローニングを行った：(1) Eco RI - Bam HI で切断されたベクター-M13mp18 [Norrrander, J., Kämpe, T., Messing, J. (1983). Construction of improved M13 vectors us-

ing oligodeoxy-nucleotide directed mutagenesis (オリゴデオキシヌクレオチド指向性突異誘発を用いる改良 M13 ベクターの組立て). Nucleic Acids Res. 21: 101-106] への Eco RI - Bgl II, (2) Xba I - Pst I で切断されたベクター-M13mp18 への Xba I - Pst I, (3) Sal I - Bam HI で切断されたベクター-M13mp18 への Xba I - Bgl II, (4) Eco RI - Pst I で切断されたベクター-M13mp19 [Norrranderら. (1983)] への Eco RI - Pst I。このため、Eco RI - Xba I 断片に含まれる Eco RI - Xba I 断片の完全ヌクレオチド配列は、Sanger, F., Nicklen, S., Coulson, A. R. [(1977). DNA sequencing with chain terminating inhibitors (鎖終結阻害剤による DNA 配列決定). Proc. Natl. Acad. Sci. USA. 74: 5463-5467] の方法により 2 本鎖で決定できる。 gdhA 遺伝子を含む Nhe I - Bgl II 断片の完全配列は 図 17 で示される (配列番号 No. 3)。

gdhA 遺伝子のヌクレオチド配列の分析

Nhe I - Bgl II 断片のヌクレオチド配列の分析から下記要素を確認することができる：

a) プロモーター (ヌクレオチド 1 ~ 572)

gdhA 遺伝子のプロモーターはそれが下記構造要素を含むことで特徴付けられる：

- ヌクレオチド 251 ~ 266

o 60 因子 [Herrick, H. J. (1983). Nitrogen control of the nif regulon in Klebsiella pneumoniae: involvement of the nifA gene and analogies between nifC and nifA (肺炎杆菌における nifA レギュロンの窒素コントロール: nifA 遺伝子の関連性と nifC 及び nifA 間の相似性). EXBIO J. 2: 39-44] により認識され及びアンモニウムにより調節されるプロモーターに特徴的な配列 TGG(Pu)A(Pu)NNNNTTGCA と類似性を示すシグナル TG GTCATATCTGTGCGG。

- ヌクレオチド 437 ~ 442

ストレプトミセス属のプロモーターの -35 領域に特徴的な配列 TTGAC(Pu) と類似性を示すシグナル TTCACA (Strohl, V. I. (1992). Compilation and analysis of DNA sequences associated with apparent streptomyces promoters (見掛け上ストレプトミセス科のプロモーターに関する DNA 配列の編集及び分析). Nucleic Acids Res. 20: 961-974)

- ヌクレオチド 466 ~ 471

ストレプトミセス属のプロモーターの -10 領域に特徴的な配列 TAG(Pu)(Pu)T と類似性を示すシグナル TAGGAT (Strohl, 1992)

- ヌクレオチド 558 ~ 572

ストレプトミセス属におけるリボソーム結合配列 AAAGCA CGTGATC と類似性を示すシグナル CGCAAACGAGGAAATC (Stro-

hi.1992)

b) コード配列 (スクレオチド 573 ~ 1913)

573 ~ 1913位にわたる既取扱は下記データからみてグルタミン酸デヒドロゲナーゼのそれに相当する：

・この既取扱から求められたタンパク質は447アミノ酸を含み、予想分子量48957ドルトンである。この分子量はC.メラッセコラの株ATCC17965のグルタミン酸デヒドロゲナーゼ同型物の変性グル電気泳動後に観察されるポリペプチドの場合(48300D)と非常に近い。

・C.メラッセコラのgdhA遺伝子のスクレオチド配列から求められたグルタミン酸デヒドロゲナーゼの一次構造は、他の生物由来のグルタミン酸デヒドロゲナーゼの一次構造と強い類似性を有する(Teller,J.X.,Smit h,R.J.,McPherson,M.J.,Engel,P.C.,Guest,J.R.(1992).The glutamate dehydrogenase gene of Clostridium sibiosum: cloning by polymerase chain reaction, sequence analysis and over-expression in escherichia c oll(クロストリジウム・シムボシウムのグルタミン酸デヒドロゲナーゼ遺伝子：ポリメラーゼ鎖反応によるクローニング、配列分析及び大腸菌内過剰発現).Eur.J.Bioche.206:151-159)。

・グルタミン酸デヒドロゲナーゼ活性にとり必須であるとしてBaker,P.J.,Britton,K.L.,Engel,P.C.,Farrant

特表平6-502548 (16)

s,G.V.,Lilley,K.S.,Rice,D.Y.,Stillman,T.J.((1992))

.Subunit assembly and active site location in the structure of glutamate dehydrogenase (グルタミン酸デヒドロゲナーゼの構造におけるサブユニットアセンブリー及び活性部位位置).Proteins.12:75-81]により述べられたアミノ酸は、C.メラッセコラのグルタミン酸デヒドロゲナーゼ中に存在し、これは Bakerら(1992)により記載された場合に相当する位置である。

・前記一次配列から求められるC.メラッセコラのグルタミン酸デヒドロゲナーゼの二次構造は他の生物のグルタミン酸デヒドロゲナーゼの二次構造と強い類似性を示す(Tellerら,1992)。

c) ターミネーター (スクレオチド 1937 ~ 1977)

gdhA遺伝子のターミネーターはそれが下記構造要素を含むことで特徴付けられる：

AG- - I3, 6 kcal/molt G Cペアリングに富むヘアピン構造を形成できる配列CCCTGATCCGCCTTAAGCTCAG CGとその後のTに富む配列TTATTGATTTCCTT。このような構造はp非依存性ターミネーターに特徴的である(Rosenberg,K.,Court,D.(1979).Regulatory sequences involved in the promotion and termination of RNA transcription(RNA転写の促進及び終結に関与する調節配列).Ann.Rev.Genet.13:319-353)。

C.メラッセコラのgdhA遺伝子の発現の調節

C.メラッセコラATCC17965のgdhA遺伝子の発現の調節はこの株が培養される培地の性質の関数としてグルタミン酸デヒドロゲナーゼ特異活性の変動を測定することにより研究された。グルタミン酸デヒドロゲナーゼ活性は超音波処理で得られたC.メラッセコラの無細胞抽出物からHeersら(1970)の方法により測定した。

この研究に用いられた培地は、ベースがLiebel,Y.,Klaeser,L.,Schleifer,K.H.((1989).Requirement of chelating compounds for the growth of Corynebacterium glutamicum in synthetic media(合成培地中コリネバクテリウム・グルタミクムの増殖に関するキレート化合物の要求).Appl.Microbiol.Biotechnol.32:205-210)により記載されたものである合成培地である。下記改変を加えた：

・炭素源は、最終1.1g/lでグルコース(培地1,2及び4)であるか又は1.0g/lでフルクトース(培地3)である。

・NH₄⁺イオンの濃度は培地1,3及び4で1.25 mMである。それは培地2で1.25 mMである(NH₄⁺を制限する)。

・培地4は最終5.0g/lのL-グルタミン酸を含有する。

前記の異なる培地で培養されたC.メラッセコラ株ATCC17965のグルタミン酸デヒドロゲナーゼに關して測定された特異活性は下記表に示される。その活性は形質転換されたNADPH2 μmol/min/mgタンパク質で表されている。

| 培地 | 培地1 | 培地2 | 培地3 | 培地4 |
|------|-----------|------------|------------|-----------|
| gdhA | 4.4+/-0.3 | 23.2+/-1.1 | 18.2+/-1.8 | 2.8+/-0.2 |

この表によってC.メラッセコラATCC17965のgdhA遺伝子の発現調節の下記3タイプを確認することができる。

・グルタミン酸による発現の抑制(倍率1.57)

・過剰アンモニウムによる発現の抑制(倍率5.27)

・グルコースによる異化抑制(フルクトースとグルコースとで倍率4.13)。異化抑制のケースにおいて、イソクエン酸デヒドロゲナーゼ、アコンターゼ及びクエン酸シンターゼ群異常活性も影響をうけることに留意すべきである。

gdhA-lacZ融合ベクターの構造

グルタミン酸、過剰アンモニウム及びグルコースによるgdhA遺伝子の調節の転写特性をC.メラッセコラで調節して、これらの調節に付されないC.メラッセコラの変異株を簡単に選択しうる手段を有するために、gdhA遺伝子の翻訳開始用プロモーター及びATGコード

シと、最初のアミノ酸が lacZ 遺伝子の領域から消失された大腸菌の lacZ オペロンとの構成体を作製した。この融合は下記のように行なった：

- gdhA 遺伝子のプロモーターを含む EcoRI - BspHI 断片の单離

- BspHI 末端からプラント末端への変換

- EcoRI 及び SmaI で置換化されたベクター pMC14-03 (Casadaban, M.J., Chou, J., Cohen, S.N. (1980). In vitro gene fusions that join an enzymatically active β -galactosidase segment to amino-terminal fragments of exogenous proteins: Escherichia coli plasmid vectors for the detection and cloning of translational initiation signals (酵素活性 β -ガラクトシダーゼ活性と外因性タンパク質のアミノ末端断片に結合させるインビトロ遺伝子融合：翻訳開始シグナルの検出及びクローニングに関する大腸菌プラスミドベクター) . J. Bacteriol. 143:971-980) に組込まれて得られた断片のクローニングでプラスミド pCGL14-3 を得る

- 前記 gdhA プロモーター - lacZ 融合体を含む pCGL14-3 の NheI - SalI 断片の单離と SpeI 及び SalI で置換化されたベクター pCGL24-1 (Reyes, O., Guyonvach, A., Bonamy, C., Saltit, V., David, F., Leblon, G. (1991). "Integron" bearing vect

ors: a method suitable for stable chromosomal integration in highly restrictive Corynebacteria ("インテグラロン" - 保持ベクター：高制限コリネバクテリアにおける安定的染色体組込みに適した方法). Gene. 107: 81-88) に組込んだクローニング、こうしてプラスミド pCGL14-0 を得る (図 1B)

- gdhA - lacZ 融合体とカナマイシン耐性を付与する λ phage 遺伝子を含む pCGL14-0 から单離されたインテグラロンのベクター pCGL12-5 への導入によって pCGL14-1 及び pCGL14-2 を得る (図 1B)。プラスミド pCGL14-1 及び pCGL14-2 を形質転換で C. メラッセコラ株 ATCC17965 に導入した。gdhA - lacZ 融合体の機能は pCGL14-1 及び pCGL14-2 で形質転換された C. メラッセコラの株 ATCC17965 における β -ガラクトシダーゼ活性、即ち pCGL12-5 で形質転換された同株に存在しない活性の検出により示された。 β -ガラクトシダーゼ活性は免色原基質 X-gal (5-ブロモ-4-クロロ-3-インドリル β -D-ガラクトビラノシド) を含有する完全固体培地 (BHI, ジフコ(Difco)) で細胞を培養することにより検出される。 β -ガラクトシダーゼ活性を有する細胞に由来するコロニーはこのような培地で青色になる。最終 1.5 g/l まで寒天を加えることで固体化され、最終 2.5 g/l までカナマイシン及び

最終 1.00 mg/l まで X-gal で補充された前記培地 1, 2, 3 及び 4において pCGL14-1 及び pCGL14-2 で形質転換された細胞を培養することにより、我々はこれら異なる培地で得られる細胞コロニーが酵素測定で示される調節と適合する着色勾配を有することから gdhA 遺伝子の調節が実際に転写タイプであることを示すことができた。実際に、培地 4 で得られるコロニーは強度増加順に培地 1, 3 及び 2 で得られる場合よりも深い青色である。我々はこの差異が培地 1 及び培地 4において pCGL14-1 で形質転換された C. メラッセコラの培養物に関する β -ガラクトシダーゼ活性の酵素測定レベルで反映されることを示した (グルタミン酸による抑制)。

| 培地 | 培地 1 | 培地 4 |
|------------------|-------|-------|
| β -gal特異活性 | 0.118 | 0.052 |

β -ガラクトシダーゼ活性は C. メラッセコラの無細胞抽出物から Miller, J.H. (1972) (Experiments in molecular genetics (分子遺伝学における実験), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York) により記載されたように測定した。

異化抑制で消失した変異株の選択

C. メラッセコラ ATCC17965 に由来する株の NTG 変異誘発を実施した。この変異誘発から、第一の選択をグルタミン酸アナログ、4-フルオログルタミン

酸に対する耐性の基準に基づき選用した。このアノログに耐性の変異株は非異化抑制種を含めた異なる種に属するかもしれない。実際には、このような変異株において、グルタミン酸デヒドロゲナーゼの特異活性に関する増加は細胞内グルタミン酸の過剰産生とひいては毒性アノログの希釈、しだがって耐性的現象を起こすことが予想できる。4-フルオログルタミン酸耐性変異株と一緒に分離し、細胞の組合せを pCGL14-1 での形質転換に付した。形質転換細胞を X-gal 及びカナマイシン含有固体培地 1 でプレート培養した。最も強度の青色を有する細胞コロニーを単離し、カナマイシン含有液体培地 1 で培養した。グルタミン酸デヒドロゲナーゼ活性を無細胞抽出物から測定し、 β -ガラクトシダーゼ活性を全トルエン処理細胞から測定した (Miller, 1972)。選択された変異株の 1 つに関して得られた結果は以下で表される。

| 活性 | グルタミン酸 デヒドロゲナーゼ | β -ガラク トシダーゼ |
|--------|--------------------|-----------------------|
| コントロール | 8.1 | 10.78 |
| 変異株 90 | 12.1 | 23.88 |

したがって得られた結果は、構築された手段で gdhA 遺伝子調節に関する変異株の表現型スクリーニングにより選択できることを実験に示す。単純にカナマイシン選択圧の非存在下で培養することにより、選択後に細胞

e I - N c o I 断片の形で c e l A のコード領域を含む断片を除去することができる。このステップは 3' で未確定配列を欠く c e l A 遺伝子を育することも可能にする。

H i n d I I I - E c o R I 断片の形における c e l A のクローニングを、大腸菌 T G 1 保有株を用いてプラスミド p M T L 2 3 に組み込み実施した。このプラスミドを保有する大腸菌株は実際には E G A の発現に関連する C M C + 表現型を保有する；得られた制限断片の分析は予想されたものと一致する。

p P R O K - c e l A の構造(図 20)は下記のとおりである：

クロントック・ラボラトリーズ社(Clonotech Laboratories, Inc.) (バロアルト, CA, U.S.A.) から入手できる 4.6 kb プラスミド p P R O K - 1 を用いる。

t a c プロモーターを含む大腸菌で複製できるこのプラスミド(Brossius et al. Gene, 27:191, 1984)を E c o R I - N c o I で加水分解する。

次いでこの制限中に E c o R I プラントの形で図 19 のアダプター-D G F 1 / D G F 2 を導入するが、これらのアダプターは B s t X I 部位を形成する。次いで c e l A を前記構築体から N a e I (プラント) - N c o I の形で導入する。

こうして得られたプラスミドは p P R O K - c e l A

から p C G L 1 4 1 及び p C G L 1 4 2 を除去することが非常に容易であることに留意すべきである。

実施例 13. ベプチドのクローニングを可能にするプラスミドの構造

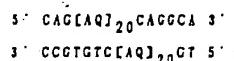
この構造のため、c e l A サブクローニングステップを実施した。プロモーター領域、その遺伝子及びもう1つの未確定遺伝子の開始部分を含む 3.5 kb H i n d I I I 断片の形で利用できる c e l A 遺伝子を、未知遺伝子断片から欠失された 2.6 kb H i n d I I I - E c o R I 断片の形で大腸菌の複数ベクター p M T L 2 3 に組み込んでサブクローニングした [Chambers, S.P., Prior, S.E., Barstov, D.A. 及び Watson, N.P. (1988). The p M T L microcloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing (p M T L - n i c) - クローニングベクター. I. ネクレオチド配列決定のために音波処理 DNA の使用を容易にする改良 p U C ポリリンカー領域]. Gene, 68:139-149]。

E c o R I 部位をその遺伝子の転写ターミネーターの直後に指向性変異挿入により導入した。この中間サブクローニングは、制限部位が導入されたとすれば、次のクローニングステップに必要である：特に、p M T L 2 3 ポリリンカーに組み込むクローニングでは E c o R I 部位直後に N c o I 制限部位を導入できるが、これは N a e I

と称される。それはアダプター-D G F 1 / D G F 2 により導入された B s t X I 部位で分離される t a c プロモーターの調節下で c e l A 遺伝子を含む。

実施例 14. マルチ A Q 配列の発現を可能にするプラスミドの組立て

20 A l a - G l u (A Q) 単位についてコードする配列の導入を実施するため、D G F 5 / D G F 6 と称される合成オリゴヌクレオチドの第二対を用いたが(図 19)、そのオリゴヌクレオチドは下記合成遺伝子に相当する：



(A Q) は A l a - G l u についてコードする配列を表す。

D G F 5 及び D G F 6 配列の末端は B s t X I 部位と適合し、したがってその配列はこの部位でクローニングである。

D G F 5 及び D G F 6 末端の配列は、一方でそれらがそのクローニング方向に向き、他方でそれらがクローニング後に B s t X I 部位を壊すような配列である。

非リン酸化アダプターの使用によれば遺列ないくつかの合成遺伝子の導入をかけることができる。

B s t X I による p P R O K - c e l A (図 20) の切断及び合成遺伝子の結合後、図 20 で表される構造を有する p P R O K (A Q) 20 c e l A が得られる。

図 21 は A Q / E G A 融合部位の構造を更に詳細に表

し、用いられる B s t X I 部位の重要性を示す。この部位の構造は以下である：



それは A T G 開始コドンとアラニンコードコドン G C A 及び決められたコード配列の後にメチオニン導入用の第二 A T G コドンを含むことが観察できる。

アダプター-D G F 5 / D G F 6 の導入は一方のみで起き、関心ある対象にとり外来である塩基を導入しない。

このプラスミドを B s t X I で処理し、同酵素で処理されたプラスミド p C G L 1 2 5 の制限産物との結合により処理する。プラスミド p C G L 1 2 5 (図 22) は複数起承 p B L 1 を含むプレビバクテリウム・ラクトファーメンタム 1 5 の複数的プラスミドである。

本発明による株はプラスミド(p C G L 1 2 5 - (A Q) 20 - c e l A) (p C G L 1 0 0 2, 図 11) での上記株の形質転換及び形質転換株の選択により得られる。

実施されたすべての融合において、標記は (A Q) 20 以前のメチオニンで始まる； (A Q) 20 と C O O H 末端でメチオニンと隣接させる予防措置も払われる；タンパク質 c e l A と融合された又はそうでないポリペプチド A Q の検出は融合タンパク質の部分的精製と臭化シアン加水分解又はその後の特異性抗体により又は分析検出により実施できる。反復ベプチドの特異的性質は容易に分離を可能にする。

當及された株は下記起原である：

大腸菌

- CLB207 reca B. バックマン (B.Bachman)
- DH5α ギブコ B R L (Gibco BRL)
- GN2929 B. バックマン
- TCI パスツール研究所 (Institut Pasteur)

プレビバクテリウム・フラブム

- ATCC 14067 ATCC

コリネバクテリウム・グルタミクム (プレビバクテリウム・ラクトファーメンタム)

- 15 S. ボナシー (S.Bonassie)
- ATCC 21088 ATCC

コリネバクテリウム・グルタミクム (コリネバクテリウム・メラッセコラ)

- ATCC 17985 ATCC

D B 5 α株はクロンティック・ラボラトリーズ (the clonal laboratories) のカタログ No C 1 0 2 1 - 1 (Palo Alto, CA, USA) から入手できる。

ATCC株は12301パークローン・ドライブ, ロックビル, MD 20852, USAのアメリカン・タイブ・カルチャー・コレクション c/o セールズ・アンド・マーケティング部門 (American Type Culture Collection c/o Sales and Marketing Department) から入手できる。

配列番号：1

配列の型：ヌクレオチドと対応タンパク質

配列の長さ：2547塩基対

塩の数：5' - 3' 方向で一本鎖表示された二本鎖

トポロジー：直鎖状

配列の種類：ゲノムDNA

起源

生物名：コリネバクテリウム・メラッセコラ (Corynebacterium melassecola)

株名：ATCC 17965

互換の実験源：クローン pCSPIC

配列の特徴

239-244 TACATA (signal-35) (S)

269-274 TAAGAT (signal-10) (S)

405-415 GACAAGCAAA リボソーム結合部位 (S)

420-2390 コード配列 (S)

420-548 分泌タンパク質ペプチド (S)

2455-2508 ヘアピン構造の依存性ターミネーター

シグナル (S)

関連生物活性：コリネバクテリウム・メラッセコラ及び
プレビバクテリウム・ラクトファーメンタムの細胞外
タンパク質PS1の前駆体

ミコバクテリウムの細胞外抗原85複合体のタンパク質
の前駆体のホモログ

株を1991年7月23日付でパスツール研究所 (パリ) のコレクション・ナショナル・デ・カルチャーズ・デ・マイクロオーガニズムス (Collection Nationale de Cultures de Microorganismes) (CNCM) に寄託した。
• No I - 1126としてプレビバクテリウム・ラクトファーメンタム 15 (CCL 2005 (B115))

| | |
|--|-----|
| AAGCTTCAGGGGAAAGCAAGGGGCTTAAAGTTATCCACAGATCCGAAAGTC | 52 |
| ATCCGGCACTGGGTGAAAGTTATCCACAGAAGGGAGGGCGGAATTGA | 104 |
| AAATTCAGCGAAATGCAGAAAGGTGGAGGGGAAATGCTGGAGCTTTCGGCG | 156 |
| ATTCGGCGGTGGCATGGAAANAGCTTAAGTTGAACTTAAGATTGAGGTC | 208 |
| ATTCGAGTTGACCTGCATCAGAACAGTTACATACCCACATATGTAAACC | 250 |
| TTCGGACTAAAGATCACCAACAGACTGAAAGAACCTGAAAGACTCTCAGGCAT | 312 |
| AGCCCACGTGTGTTGTCGGGCGGAAGCGGGGAACTTCGGGACGGCTA | 364 |
| ACTCATGGGGGCTGTGGCAGTATCCAAAGATCAGAAAGCAGAAC | 416 |
| TTC ATG CGC GAC ACC GCA TTG CGT TCC ATC AAG CCT AAA Met Arg Asp Thr Ala Phe Arg Ser Ile Lys Ala Lys | 455 |
| GCT CAG CCT AAG CGC CGT TCC CTC TGG ATT GCA GCA GGC Ala Glu Ala Lys Arg Ser Leu Thr Ile Ala Ala Gly | 494 |
| GCT GTC CCA ACC GCA ATT CGG TTG ACT ATG TCC CTG GCA Ala Val Pro Thr Ala Ile Ala Leu Thr Met Ser Leu Ala | 533 |
| CCT ATG CCT TCG CCT CAA TCC ACC AAC CTT TCC TCT GAT Pro Met Ala Ser Ala Glu Ser Asn Leu Ser Asp | 572 |
| GCC GTC GTC ACC ATC CGG CGC CGC GTC ACC CAT CGC Ala Val Val Gly Ser Ile Ala Glu Gly Val Thr Asp Gly | 611 |
| CTG ACT GAC TAC TCG AAC CCT CGC GTC CGG CCT CCT Leu Thr Asp Tyr Leu Lys Pro Arg Val Glu Glu Leu Pro | 650 |
| GCT GGT GAA GTC ACC TAC CCA GAG ATC CCC CGC CTG CCT Ala Glu Val Thr Tyr Pro Glu Ile Ala Glu Leu Pro | 689 |
| GAT GGT GTG CGC GTG ATC ACC GCT CGG TGG GCA ACC TCC Asp Glu Val Arg Val Ile Ser Ala Glu Thr Ile Ser | 728 |
| AAG CAT GTC ATT TTG ACT ATT CAG TCT GCA GCA ATG CCA Lys His Val Thr Ile Leu Thr Ile Glu Ser Ala Ala Met Pro | 767 |
| GAG CCC CCA ATC AAG GTG CAG CTG CTG CTT CGG CGT GAC Glu Arg Pro Ile Lys Val Glu Leu Leu Pro Arg Asp | 806 |
| TGG TAC TCT TCC CGG AAC CGT GAG TTC CCT GAA ATC TGG Ile Tyr Ser Ser Pro Asn Arg Glu Phe Pro Glu Ile Thr | 845 |
| GCA CCT GAC CGT CTG CGC CGG ATT GAA GAG CGG AGT GGT Ala Leu Asp Gly Leu Arg Ala Ile Glu Glu Glu Ser Glu | 884 |

特表平6-502548 (20)

| | | | |
|---|------|---|------|
| TGG ACC ATT GAG ACC AAC ATT GAG CAG TAC TAC GCC GAT | 923 | CAG GCG TTC CCT CAC ATC GCT AAC GCT CTT GCC ATG TCC | 1586 |
| Trp Thr Ile Glu Thr Asn Ile Glu Glu Tyr Tyr Ala Asp | | Gln Ala Phe Pro His Ile Ala Asn Ala Leu Gly Met Ser | |
| AAG AAC GGC ATT GTT CGC CTC CCA AIC GGT GGC GAG AGC | 962 | ACT GAG GAC CGT GGC GTT GAG TGT GCA CCT GTC GGC GCA | 1623 |
| Lys Asn Ala Ile Val Val Leu Pro Ile Gly Glu Ser | | The Glu Asp Arg Gly Val Glu Cys Ala Pro Val Gly Ala | |
| TCC TTC TAC TCT GAC TGG GAA GAG CCA AAC AAC GGC AAG | 1001 | ATC GCT GAC GCT GTT GGC GAC GGC GCG ATC GGC ACC TCC | 1664 |
| Ser Phe Tyr Ser Asp Trp Glu Glu Pro Asn Asn Gly Lys | | Ile Ala Asp Ala Val Ala Asp Gly Ala Met Gly The Cys | |
| AAC TAC CGG TGG GAG ACC TTC CTG ACT CAG GAG CTC GCA | 1040 | CTG ACC AAC GAA TAC GAT GTT ACC GGC GGT AAG GGC CAG | 1703 |
| Asn Tyr Glu Trp Glu Thr Phe Leu Thr Glu Leu Ala | | Leu Thr Asn Glu Tyr Asp Val The Gly Gly Lys Ala Glu | |
| CGG ATC CTG GAC AAG GGC TTC CGT TCC AAC ACC GAT GCG | 1079 | GAC TTG CCT AAC GGT CGC GCA TAC TGG TCT GCA AAC ACT | 1742 |
| Pro Ile Leu Asp Lys Gly Phe Arg Ser Asn Thr Asp Arg | | Asp Phe Ala Asn Gly Arg Ala Thr Trp Ser Ala Asn Thr | |
| GCC ATC ACC GGT ATC TCC ATG GGC GGT ACC GCT GCG GTT | 1118 | GCC CCT TTC CGC GTT GGA CGC ATC AAC GCT CGT TAC TAC | 1781 |
| Ala Ile Thr Gly Ile Ser Met Gly Lys Thr Ala Ala Val | | Gly Ala Phe Gly Arg Leu Val Gly Arg Ile Asn Ala Arg Tyr | |
| AAC ATC GCA ACC CAC CCA GAC ATG TTT AAG TTC GTC | 1157 | TCT GAG CGT GGT GGA CGC TAC TGG TCT GGC TAC CCA | 1820 |
| Asn Ile Ala Thr His His Pro Asp Met Phe Lys Phe Val | | Ser Glu Leu Gly Pro Asp Ser Trp Leu Gly Tyr Pro | |
| GGT TCC TTC TCC GGC TAT CTG GAC ACC ACC TCC GCT GGC | 1196 | ACC TCT TCT GAG TTG AAG ACA CCA GAC GGA CGT GGC CCC | 1859 |
| Gly Ser Phe Ser Gly Tyr Leu Asp Thr Ser Ala Gly | | The Ser Ser Glu Leu Lys Thr Pro Asp Gly Arg Gly Arg | |
| ATG CCA ATC GCT ATT TCC GCA GCC CTG GCA GAC GGC GGC | 1235 | TTC GTC ACC TTC GAG CAC GGC TCC ATC TAC TGG ACC GGC | 1898 |
| Met Pro Ile Ala Ile Ser Ala Leu Ala Asp Gly | | Phe Val Thr Phe Glu His Gly Ser Ile Tyr Trp Thr Ala | |
| GGA TAC GAT GCC AAC GCA ATG TGG GGA CCA GTC GGT TCT | 1274 | ACC ACT GGT CCT TGG GAA ATC CCA GGC TAC ATG CTC CCC | 1937 |
| Gly Tyr Asp Ala Asn Ala Met Trp Gly Pro Val Gly Ser | | The Ser Ser Glu Leu Lys Thr Pro Asp Gly Arg Gly Leu | |
| GAC CCC TGG CAG GAA AAC GAT CCA AAC AGC AAC GTA GAC | 1313 | GCA TGG GGC ACC CAG GAC TAT GAG AAC GGC ACC CTC GGC | 1976 |
| Glu Arg Trp Glu Asn Asp Pro Lys Ser Asn Val Asp | | Ala Trp Gly Thr Glu Asp Tyr Glu Lys Gly Ser Leu Gly | |
| AAG CTC AAC GGC AAG ACC ATC TAC GTT TCC TCT GGT AAC | 1352 | TAC CCA ACC GGC GGC GCA GTT GAA TAC AAC GGT GGC CGC | 2015 |
| Lys Leu Lys Gly Lys Thr Ile Tyr Val Ser Ser Gly Asn | | Tyr Pro Thr Gly Ala Ala Val Glu Tyr Asn Gly Glu Leu | |
| GGT GCA GAT GAC TTC GGT AAG GAA GAC GCT GTC GCT ATT | 1391 | CGC CAG CAG TTC GAA GGT GGC TAC GCA TCC CGT ACC TCC | 2054 |
| Gly Ala Asp Asp Phe Gly Lys Glu Asp Ser Val Ala Ile | | Arg Gln Gln Phe Glu Gly Tyr Val Phe Arg Thr Ser | |
| GGA CCT GCA AAC GCG ACA GGT GTC GGT CGT GAA GTT ATC | 1430 | ATC AAC CAG TCT TAC TGG GTT CGC GGA GAA ATC TCC ARG | 2093 |
| Gly Pro Ala Asn Ala Thr Gly Val Glu Leu Glu Val Ile | | Asn Asn Gln Ser Tyr Trp Val Arg Gly Ile Ser Lys | |
| TCC CGT ATG ACT TCC CAG ACC TTC GTC GAT CGT GCA AAC | 1469 | ARG TAC GGC GAT GAC GGA ATC TTC GCT GAG CTT GGT TCC | 2132 |
| Ser Arg Met Thr Ser Asn Glu Thr Phe Val Asp Arg Ala Asn | | Lys Tyr Ala Asp Asp Gly Ile Phe Ala Glu Leu Gly Phe | |
| CAG GCT GGC GTG GAA GTT GTT GCT ACC TTC CGT CCA TCC | 1508 | CCA ACC GGC AAT GAG AAG TTG ATC AAC GGT GGC GCT TTC | 2171 |
| Gln Ala Val Glu Val Ala Phe Arg Pro Ser | | Phe Thr Gly Asn Glu Lys Leu Ile Asn Gly Glu Ala Phe | |
| GCC GTG CAC TCA TGG GAA TAC TGG CAG TTC GAG ATG ACT | 1547 | CAG GAA TTC GAA AAG GGC AAC ATC TAC TGG TCC CGT TCC | 2210 |
| Gly Val His Ser Trp Glu Tyr Trp Glu Phe Glu Met Thr | | Gln Glu Lys Gly Asn Ile Tyr Trp Ser Val Ser | |

| | |
|---|------|
| ACT GGC GCG CAC GTG ATT CTG CAC GGC GAC ATC TTC GAC | 2249 |
| The Gly Ala His Val Ile Leu His Gly Asp Ile Phe Asp | |
| GCA TGG GGT GCT AAG GGC TGG GAG CAG GGC GAA TAC GCG | 2288 |
| Ile Trp Gly Ala Lys Gly Trp Glu Glu Gly Glu Tyr Gly | |
| TTC CCA ACC TCT GAC CAG ACC GCA ATC ACC GCG GGT GGA | 2327 |
| Thr Pro Thr Ser Asp Asn Glu Thr Ala Ile The Ala Gly Gly | |
| CAG ACC ATT GAT TTC CAG AAC GGC ACC ATC CGT CAG GTC | 2366 |
| Gin Thr Ile Asp Phe Glu Asn Gly The Ile Arg Glu Val | |
| AAT GGC CGA ATT GAG GAG TCT CGC TAATAGTGAA AGCCCCAATCA | 2409 |
| Asn Gly Arg Ile Glu Glu Ser Arg | |
| CGCAACTCTCGCTTCCGGACTTTTGCTGCCCTGAGCTTGCTTGGGGGGAA | 2451 |
| GTCAGCTGTTGAGGGAGATGATTCTCCCTCGAACAGGGCCAGCCCCAACAGAAA | 2513 |
| CGAGCCGCTGGGTCAAGCAGCACCCCAAGGTGGC | 2547 |

配列番号：2

配列の型：スクレオチドと対応タンパク質

配列の長さ：2702塩基対

鎖の数：5' - 3' 方向で一本鎖表示された二本鎖

トポロジー：直鎖状

配列の種類：ゲノムDNA

起源

生物名：コリネバクテリウム・メラッセコラ (*Corynebacterium selassescolum*)

株名：ATCC 17965

直接の実験源：クローンpCGL815、pCGL824

配列の特徴

S62-S67 AACGAGリボソーム結合部位(S)

S79-2108 コード配列(S)

S79-888 分泌タンパク質シグナルペプチド(S)

2188-2333 ヘアピン構造転写ターミネーター・シグナル(S)

関連生物活性：コリネバクテリウム・メラッセコラ及び
ブレビバクテリウム・ラクトファーメンタムの壁の
外層表面層を構成するタンパク質と細胞外抗原PS2
の前駆体

附表平6-502548 (21)

GAATCCCTGGAATTAGCCGGTTTAGTACTTTCAAGGGTGCTCATTTCTAC 52
 CAGATCTCACTTGTGGGTAGAGTCACCTGAAATAATTGCAACCGCACCG 104
 GTGATATGCTTATTCGCTAAGTAGTTCGAGGTTAACGTGATTTAGGT 156
 AACAAATTTCAGCTCGGGTAAAGACTTCTATGCCCTCAGAGCTCTAT 208
 TAGGAAACTGACACCCTGATTAATAGCCYACCCCCGAAITGGGGATG 260
 GGTCAITTTGCTGTGAAGGTAGTTTGATGCAATAGACCTGGCTTATAA 312
 AGAAAATGAAACGTCAGATCGATATAAAAAGAACAGTTGACTTCAGGT 364
 TTGAGCATTTCCTCCGATTCGGCTGGCAAAATCTAATTGCTCTTACAG 416
 TTTTCTCACGACAGGCTGCTAAAGCTGCTAGTTCGGGCCAGTGACTGG 468
 CGTTTACTGGATATAAAAGTAAATCCCATGTCGATCAGCCATTGGGTGT 520
 TTCCATAGCAATCCAAGGTTGCTTCTGGATCACCTATTCAGGAGCTTC 572
 GCCTCT ATG TTT AAC AAC CGT ATC CGC ACT GCA GCT CTT 611
 Met Phe Asn Asn Arg Ile Arg Thr Ala Ala Leu
 GCT GGT GCA ATC GCA ATC TCC ACC GCA GCT TCC CGC GTC 650
 Ala Gly Ala Ile Ala Ile Ser Thr Ala Ala Ser Gly Val
 GCT ATC CCA GCA TTC GCT CAG CAG ACC GCA ACT TTC 689
 Ala Ile Pro Ala Phe Ala Glu Thr Asn Pro Thr Phe
 AAC ATC ACC AAC GGC TTC AAC GAT GCT GAT GCA TCC ACC 728
 Asn Ile Thr Asn Gly Phe Asn Asp Ala Asp Gly Ser Thr
 ATC CAG CCA GTT GGC CCT GTT AAC CAC ACC GAG GAA ACC 767
 Ile Glu Pro Val Gly Val Asn His Thr Glu Glu Ala
 CTC CGC GAC CTG ACT GAC TCC ACC GGC GCT TAC CTG GAA 806
 Leu Arg Asp Leu Thr Asp Ser Thr Gly Ala Tyr Leu Glu
 GAG TTC CAG AAC GGC ACC GTC CAG GAA ATC GTC GCA GCA 845
 Phe Asn Glu Asn Gly Thr Val Glu Glu Ile Val Glu Ala
 TAC CTG CAG GTT CAG GCT TCC GCA GAC GGA TTC GAT CCT 884
 Tyr Leu Glu Val Glu Ala Ser Ala Asp Gly Phe Asp Pro
 TCT GAG CAG GCT GCT TAC GAG GCT TTC GAG GCT GCT GCG 923
 Ser Glu Ala Ala Ala Tyr Glu Ala Phe Glu Ala Ala Arg
 GTC CGT GCA TCC CAG GAG CTC GCA CCT TCC CCT GAC ACC 962
 Val Arg Ala Ser Glu Glu Leu Ala Ile Ser Ala Glu Thr

CAG CAC TAC TCT GGG CTG ATC CCT AAC GTC TTC ATC GCA 1625
 Glu His Tyr Ser Ala Leu Ile Val Asn Leu Phe Ile Ala
 GCA GTT GCA AAC ATC AGC GAG CTC AAC GCT GCA GAT GCT 1664
 Ala Val Ala Asn Ile Ser Glu Leu Asn Ala Ala Asp Ala
 GAA GCA GCA GCT TAC TAC CTC CAC TGG GAC ACC GAC CTC 1703
 Glu Ala Ala Ala Tyr Tyr Leu His Trp Asp Thr Asp Leu
 GCA ACC AAC GAT CAG GAC GAA GCT TAC TAC AAG GCT AAG 1742
 Ala Thr Asn Asp Glu Asp Glu Ala Tyr Tyr Ala Lys
 CTC GAC TTC GCT ATC GAG ACC TAC GCA AAG ATC CTG TTC 1781
 Leu Asp Phe Ala Ile Glu Thr Ala Lys Ile Leu Phe
 AAC CGT GAA GTT TGG CAG CAG CCA CTG CCT TAC GTC CAG 1820
 Asn Glu Val Trp Glu Glu Pro Leu Ala Tyr Val Glu
 AAC CTG GAT GCA GGC GCA CGT CAG GAA GCA GCT GAC CCT 1859
 Asn Leu Asp Ala Gly, Ala Arg Glu Glu Ala Ala Asp Arg
 GAG GCA CCT CGC GCA GCT GAC GAA GCT TAC CGC GCT GAG 1898
 Glu Ala Ala Arg Ala Ala Asp Glu Ala Tyr Arg Ala Glu
 CAG CTC CGC ATC CCT CAG GAA GCA GCT GAC CCT CAG AAC 1937
 Glu Leu Arg Ile Ala Glu Glu Ala Ala Asp Ala Glu Lys
 GCT ATC GCT GAG GCG CCT CCT AAC GAA GCA GAA GGC AAC 1976
 Ala Ile Ala Glu Ala Leu Ala Lys Glu Ala Glu Glu Asn
 AAC GAC AAC TCC TCC GAC AAC ACC GAG ACC GGT TCT TCT 2015
 Asn Asp Asn Ser Ser Asp Asn The Glu Thr Gly Ser Ser
 GAC ATC GGA TCC TGG GGA CCT TTC GCA GCA ATT GCA GCT 2054
 Asp Ile Glu Ser Trp Glu Pro Phe Ala Ala Ile Ala Ala
 ATC ATC GCA GCA ATC GCA GCT ATC TTC GCA TTC CTC TCC 2093
 Ile Ile Ala Ala Ile Ala Ala Ile Phe Pro Phe Leu Ser
 GGT ATC GTT AAC TTC TAA TTTCGAACCCGAGATAGCTAAAGTTAAA 2139
 Glu Ile Val Lys Phe
 CCACCTCCCTTCTGGGGAGGTGGTTTCTGGCTAACAGCACCAAAA 2191
 GAAAAGCCACCTCTTGTATCTCAAGGAGCTGGCTTACTTTTATTACTGGG 2243
 GAGCGGGAGGTGGCTGCTGATAGCAAAATCTTTGCTTTTACGGAACT 2295

ATC ACC AAC ACC CGC GAG TCC GCT GCT TAC GCA CTC AAC 1001
 Ile Thr Lys Thr Arg Glu Ser Val Ala Tyr Ala Leu Lys
 GTT GAC CAG GAA GCT ACC GCT GCT TTC GAG GCA TAC CCG 1040
 Val Asp Glu Glu Ala Thr Ala Ala Phe Glu Ala Ala Tyr Arg
 AAC GCA CTT CGC GAT GCA GCT ATC TCT ATC AAC CCA GAT 1079
 Asn Ala Leu Arg Asp Ala Ala Ile Ser Ile Asn Pro Asp
 GGC TCT ATC AAC CCA GAT ACC TCT ATC AAC GCA CTG ATC 1118
 Gly Ser Ile Asn Pro Asp Thr Ser Ile Asn Leu Ile
 GAT GCT GCT AAC GCT GCT AAC CGC ACC GAT GCT GCA GAG 1157
 Asp Ala Ala Asn Ala Ala Asn Arg Thr Asp Arg Ala Glu
 ATC GAG GAT TAC GCT CAC CTT TAC ACC CAG ACC GAT ATT 1196
 Ile Glu Asp Tyr Ala His Leu Tyr The Glu Thr Asp Ile
 GCT CTT GAA ACT CCA CAG CCT GCA TAC GCT TTC GCA GAG 1235
 Ala Leu Glu The Pro Glu Leu Ala Tyr Ala Phe Glu Asp
 CTG AAG GCT CTT CAG GCT GAG GTC GAC GCA GAC TTC GAG 1274
 Leu Lys Ala Leu Glu Ala Val Asp Ala Asp Phe Glu
 TGG TTG GGC GAG TTC GGA ATC GAC CAG GAA GAC GGT AAC 1313
 Ile Trp Leu Gly Glu Phe Gly Ile Asp Glu Asp Gly Asn
 TAC GTT CAG CGC TAC CAC CCT CCT CCT GCA GAG GCA CTC 1352
 Tyr Val Glu Arg Tyr His Leu Pro Ala Val Glu Ala Leu
 AAG GCT GAG GTC GAC GCT CGC GTC GCA GCA ATT GAG CCA 1391
 Lys Ala Glu Val Asp Ala Arg Val Ala Ala Ile Glu Pro
 CTT CGT GCA GAC TCC ATC GCT AAC CCT GAG GCG CAG 1430
 Leu Arg Ala Asp Ser Ile Ala Lys Asn Leu Glu Ala Glu
 AAG TCT GAC GTT CTG GTT CGC CAG CTC TTC CTC GAG CCT 1469
 Lys Ser Asp Val Leu Val Arg Glu Leu Phe Leu Glu Arg
 GCA ACC GCA CAG CGC GAC ACC CTC CGT GTT GCA GAG GCG 1508
 Ala Thr Ala Glu Arg Asp The Leu Arg Val Val Glu Ala
 ATC TTC TGT ACC TCT CCT CGT TAC GTT GAA CTC TAC GAG 1547
 Ile Phe Ser Thr Ser Ala Arg Tyr Val Glu Leu Tyr Glu
 AAC GTC GAG AAC GTT AAC GTT GAG AAC AAC ACC CTT CGC 1586
 Asn Val Glu Asn Val Asn Val Glu Asn Lys Thr Leu Arg

GATAATCGGCTTAATGACTTCCCACTGGGGGAACTCCGAAAGGGCATTTGA 2347
 TTGTTCCAGGGTAAAGTGGCAAGGAGCTTCCTGATCGGGAACTTGGCTG 2399
 GCCCCCACAAATGACCCAGGGGAGGGCATGAAATCTGAGGGACGGGCTGGCC 2451
 TCAATGATGGCTGGAACTTCCAAACCACGGGACCTGACGGGCCAACCTGCA 2503
 AGGTAGCTTCTGGTGGCAAGGGGAGGGGAGGGAGGAGACGGACGGTACCGGT 2555
 GATGCCCAAGGAATCGGCTGCTTGGCTGGTACGGCCACGACCCAGTGTCA 2607
 TCGAGAGGCAATTGCAACCCATCGGGGGTCACTTCTGATTITGTTCCGAG 2659
 GATCTCTATCCCTSGATGCGTGGGTAGCTCCGAGCTC 2702

持表平6-502548 (22)

配列番号 : 3
 配列の型 : ネクレオチドと対応タンパク質
 配列の長さ : 2160 基塩基
 構造の数 : 5' - 3' 方向で一本鎖表示された二本鎖
 トポロジー : 直鎖状
 配列の種類 : ゲノム D N A
 起源
 生物名 : コリネバクテリウム・メラッセコラ (*Corynebacterium melassecola*)
 株名 : ATCC 17965
 直接の実験群 : クローン pCGL315 , pCGL313 , pCGL310
 配列の特徴
 因子 60 により認識され及びアンモニウムにより調節
 される 251-268 TGCTCATATCTGTGGC プロモーター部位
 (9)
 437-442 TTCAACA プロモーターシグナル領域 - 35 (S)
 468-471 TACCAT プロモーターシグナル領域 - 10 (S)
 551-572 CCCAACCGACCAAATC リボソーム結合部位 (S)
 573-1913 コード配列 (S)
 1937-1977 ヘアピン構造の非依存性転写ターミネーター
 シグナル (S)
 関連生物活性 : 48300d ポリペプチドとして変性グルクロン酸
 活動する N A D P H 依存性グルタミン酸デヒドロゲナーゼ活性

TCC GAC TTC GAC CCT AAG GGC AAC TCC GAT CTG GAA ATC 1001
sec asp phe asp pro lys gly lys ser asp leu glu ile

ATG CGT TTC TCC CAG TCC TTC ATG ACC GAG CTG CAC CCC 1040
met cys phe asp met thet glu his ser leu phe met his

CAC ATC GGT GAG TAG CGC GAC GTT CCT GCA GGT GAC ATC 1079
his ile gly glu tyr arg asp val pro ala gly asp ile

GGA GTT GGT GCC CGC GAG ATC GGT TAC CTG TTT GGC CAC 1118
val val gly gly arg glu ile gly tyr leu phe glu his

TAC CGT CCC ATG GCC AAC CAG CAC GAG TCC GGC GTT TTG 1157
tyr arg arg met ala asn gln his glu sar gly val leu

ACC GGT AAG GGC CTG ACC TGG GGT GGA TCC CTG GTC GCG 1196
thu gly lys gly leu thr trp gly gly ser leu val arg

ACC GAG GCA ACT GGC TAC CGC TCC GTT TAC TTC GTG ACT 1235
the glu ala thet gly tyr gly cys val tyro phe val ser

GAA ATG ATC AAG CCT AAG GCC CAG AGC ATC ACC GGC CAG 1274
glu met ile lys ala lys gly glu ser ile ser gly gln

AAG ATC ATC GTC TCC GGT TCC CGC AAC GTC GCA ACC TAC 1313
ala ile ile val sar gly ser gly val val ala thet tyr

GCC ATT GAA AAG CCT CAG GAA CTC CGG GCA ACC GTT ATT 1352
ala ile glu lys ala gln glu leu gly ala thet val ile

GGT TTC TCC GAT TCC AGC GGT TGG GTT CAT ACC CCT AAT 1391
gly phe ser asp ser ser gly trp val his thet pro asn

GCC GTT GAC GTG GCT AAG CTC CGC GAA ATC AAG GAA GAT 1430
gly val asp val ala lys leu arg gly ile lys glu val

CCG CGG GCA CGC GTA TCC GTG TAC GCC GAC GAA GTT GAA 1469
arg arg ala arg val ser val tyr ala asp glu val glu

GGC GCA ACC TAC CAC ACC GAC GGG TCC ATC TGG GAT CTC 1508
ala thet thet his thet asp gly ser ile trp asp leu

AAG TGC GAT ATC GCT CCT TGT GCA ACT CAG AAC GAG 1547
lys cys asp ile ala leu pro cys ala thet gln asn glu

CTC AAC GGT GAG AAC GCT AAG ACT CTT GCA GAC AAC GGC 1586
asn gly glu asn ala lys thet leu ala asp asn gly

GGC CGT TTC GCT GGT GAA CGC AAC ATG CCT TCC ACC 1625
cys asp phe val ala glu gly ala asn met pro ser thet

CCTAGCCCTGGGAGCTCTAGGAGATTGTGAAACCGGTAAATTCTCCCA
 TGCAGGCCCTATAAAGTCTGTAACCATTCATTGGGGTCTCAAGTGTGG 104
 CCAGGTTATATAACCACTGAGTCACCTGGTCTCATTCGCTGGTCGAAT 156
 TTAACTAAAGAGAGACTCTATCAGTAACTGGCGTTGGGCAATACACAA 208
 TTGAAACCTAAAGAAATTTCAGAACATTTTAACTCTTGTTGGTCATATC 260
 TTGTCGACACTGCCATAATTGAACTGTGAGGATTACCGCTAAATGCCCGC 312
 AGTGAGTTAAAGTCTAAAGCAAGAGTTGGCTCTTAGGGCATCCGTAGTTA 364
 AAACATTAAACCGTTAGGTATGACAAGCCGGTTGATGTGAAACGGAGTTTA 416
 AAAGTTCCAGGAATCAGATTTCACAGGCAATTGGTCCAGCAAACCGCTAG 468
 GATGTACATGGTGCCTCAATGGGAAACCACCUACATCACTAAATGGCCAGA 520
 TACACACCTTAAATCTGCGCCGCACTGCGCCGAGATGGGAAACGGAGAAC 572

 ATG ACA GTT CAT GAG CAG GTC TCT AAC TAT TAC GAC ATG 611
 met the val asp glu val val asn tyr tyr asp met

 CTT CTG AAG CGC AAT GCT CCC GAG CCT GAA TTT CAC CAG 650
 leu leu lys arg asn ala gly glu pro glu phe his gln

 GCA GTG GCA CAG GTT TTG GAA TCT TTG AAG ATG CTC CTG 689
 ala val ala glu val leu glu ser leu tyr ile val leu

 GAA AAG GAC CCT CAT TAC GCT GAT TAC GGT CTC ATC CAG 728
 glu lys asp pro his tyr ala asp tyr gly leu ile gln

 CGC CTG TCG GAG CCT GAG CGT CAG CTC ATC TTC CGT GTC 767
 arg leu cys glu pro glu arg gln leu ile phe arg val

 CCT TGG GTT GAI GAC CAG CCC GAG GTC CTC GTC AAC CCT 806
 pro trp val asp asp gln gly gln val his val asn arg

 GGT TTC CCC GTG CAG TTC AAC TCT GCA CTC CGA CCA TAC 845
 gly phe arg val gln phe asn ser ala leu gly pro tyr

 AAG GCC CCC CTG CCC TTC CAC CCA TCT GTA AAC CTG GGC 884
 lys gly gly leu arg phe his pro ser val asn leu gly

 ATT GTG AAG TTC CTG GGC TTT CAG CAG ATC TCT AAA AAC 923
 ile val lys phe leu gly phe glu gln ile phe lys asn

 TCC CTA ACC GGC CTG CCA ATC GGT GGC AAC GGT GCA 962
 ser leu thr gly leu pro ile gly gly lys gly lys gly

CCA GAG CCT GTT GAG GTC TTC CGT GAG CCC GAC ATC CCC 1664
 pro glu ala val glu val phe arg glu arg asp ile arg

 TTC GCA CCA GCC AAC GCA GCT AAC GCT GGT GGC GTT GCA 1703
 phe gly phe gly lys ala ala asn ala gly gly val ala

 ACC TCC GCT CTG GAG ATC CAA GAC GTC TGG CCC GAT 1742
 thr sec ala leu glu met gln gln asn ala ser arg asp

 TCC TGG ACC TTC GAG TAC ACC GAC GAG CCA CTC CAG GTC 1781
 sec ter phe glu tyr thr asp glu arg leu glu val val

 ATC ATG AAG AAC ATC ATC TTC AAG ACC TGT GCA GAG ACC GCA 1820
 ile met lys asn ile phe lys thr cys ala glu thr ala

 GCA GAG TAT GCA CAC GAG AAC GAT TAC GTC GTC GGC GCT 1859
 ala gla tyr his glu amu tyr val val val val ala

 AAC ATT GCG GGC TTC AAG AAG GTC GAT GCG GCG ARG CTG 1898
 asn ile ala gly phe lys lys val ala asp ala met leu

 GCA CAG GGC GTC ATC TAA GACCCCTGCACTTTACTTAAACCCCTGCA 1944
 ala glu gly val ile CCA

 TCCGGCTTAAGCATCAGGGATTTTGATTCTTCTCAGGTCAATTATGCCAAC 1996
 CACATGGGTTTAAGCTCAGCTTGCGGCTGCGGCAATCATGATCACCGGGTGTCT 2048
 TTAACTGTGGCCACAGTCCTGGGAAAGATCCGCTTGTATTGACCGCATCTGGT 2100
 GCGCTGGGCTTCACTGACAACTCAGTACCTCAGGGCTGGCTGCCAAGAAC 2152

 CCCCAGGAGGACGGCTGTGCGCTGCCCTCGAGAAGGC 22190

図面中の符号図 8 :

A - p : csp1 のプロモーター
 s : 検定 csp1 シグナル配列
 m : 成熟 PS 1 の最初の 30 アミノ酸

図 9 :

A - p : csp1 のプロモーター
 s : 検定 csp1 シグナル配列
 m : 成熟 PS 1 の最初の 30 アミノ酸

5' から 3' への配列の詳細

ACACCGCATTTGTTCCATCAAGGCTAACGGCTAACGGCTAACGGCGTTCCTCCTGGATTGCAAGCAGGCCGCTGCTAACGGCAATTGGCTTGACTATGTCCCCTGCACCATATGGCTTCGGCTCAGTCAGGAACCTTTCCTCTGATGCCGTAGTTGGCACGCGCATGGCGTACCGATGGCTGACTGAGTACCTGAAAGCCTCGGGCTGAGGACCTGGCAATTCATGGCAATGCCGTOTCGACCCCGCAAACACTGTGTGAGGGCAGGGCAAGGCAATGGCTTAAACAAAATAACCCATGGCTACTCTATGGCATAATGAGTCGGAGTAAGTGAATGCTAAAGCTCAAAGCAGAATGGGAAGACTGGGAAGAGCAAGAGAAATTCTCGAACCGTGCAGGGAGGATA

図 10 :

A - p : csp1 のプロモーター
 s : 検定 csp1 シグナル配列
 m : 成熟 PS 1 の最初の 30 アミノ酸

5' から 3' への配列の詳細

CCTTCTCTGATGCCGTAGTGGCAGCATCGGCAGGGCTCACCGATGGCCTGACTGACTACCTGAAGGCTCGCTCGAAGACCTGAGGCCAAATTGGATGGCACAGAAGGACAGAGGCAAGAGGACAGAAGGACAGAAGGACAGAAGGCAATGCCGGCCTGTCGACCCGGAAACACTGTGTCAGGGCAGGTG

図 11 :

A - p : csp1 のプロモーター
 s : 検定 csp1 シグナル配列
 m : 成熟 PS 1 の最初の 30 アミノ酸

5' から 3' への配列の詳細

CCTTCTCTGATGCCGTAGTGGCAGCATCGGCAGGGCTCACCGATGGCCTGACTGACTACCTGAAGGCTCGCTCGAAGACCTGAGGCCAAATTGGATGGCACAGAAGGACAGAGGCAAGAGGACAGAAGGACAGAAGGACAGAAGGCAATGCCGGCCTGTCGACCCGGAAACACTGTGTCAGGGCAGGTG

図 12 :

ポリリンカー 1 : 0.001/SacII,BstXI,NotI,XbaI.
 ポリリンカー 2 : 1.531/ClaI,SalI,AatI,MluI,NcoI,
 BgIII,XbaI,SstI,PstI,SmaI,BamHI,SpeI.
 ポリリンカー 3 : 1.561/IbaI,NotI,SacII,BstII.

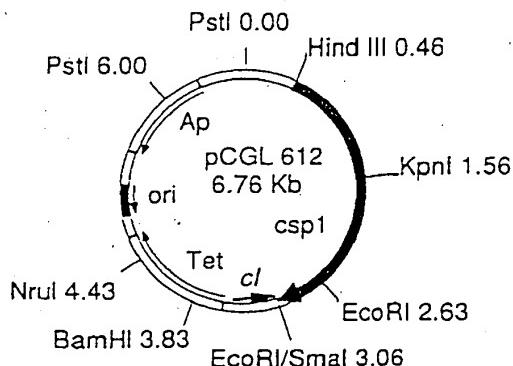


FIG. 1

| | |
|---|-----|
| AGCTTCAGGGAAAAACAGGCC | 52 |
| ATCCGGCAGCTGGGT | 104 |
| XXXXTCAGCAAATCGAA | 156 |
| ATCCCCGGCTGGCAATTCA | 208 |
| ATTCCTGAACTGTGACCTG | 260 |
| TTCTGGACTAAGATCACG | 312 |
| ACCCCCAGCTGTGTTGCG | 364 |
| ACTCATTCGGGCTGTGCG | 416 |
| TTC ATG CGC GAC ACC GCA TTT CGT TCC ATC AAC GCT AAA | 455 |
| Met Arg Asp Thr Ala Phe Arg Ser Ile Lys Ala Lys | |
| GCT CAG GCT AAC CGC CGT TCC CTC TGG ATT GCA GCA GGC | 494 |
| Ala Glu Val Lys Arg Ser Leu Tyr Ile Ala Ala Gly | |
| GCT GTC CCA ACC GCA ATT CGG TTG ACT ATG TCC CTG GCA | 533 |
| Ala Val Pro Thr Ala Ile Ala Leu Thr Met Ser Leu Ala | |
| CCT ATG CCT TCC GCT CGA TCC AGC AAC CCT TCC TCT GAT | 572 |
| Pro Met Ala Ser Ala Ser Ser Leu Ser Ser Arg | |
| GCC GTC GTT GGC AGC ATC CGG CAG GGC GTC ACC GAT GGC | 611 |
| Ala Val Val Gly Ser Ile Ala Glu Gly Val Thr Arg Gly | |
| CTG ACT GAC TAC CGT AAG CCT CGC GTC GAA GAG CCT CCT | 650 |
| Ile Thr Asp Tyr Leu Phe Arg Val Val Glu Glu Leu Pro | |
| GCT GGT GAA GTC ACC TAC CGA CGG ATC CGC GGG CGT CCT | 689 |
| Ala Glu Val Tyr Val Glu Ile Ala Glu Lys Pro | |
| GAT GGT GTC CGC CGT ATC ACC GCT GAG TCC GCA ACC TCC | 728 |
| Asp Gly Val Arg Val Ile Ser Ala Glu Thr Pro Ala Thr Ser | |
| AAG CAT GTC ATT TTG ACT ATC CGG TCC GCA GCA ATG GCA | 767 |
| Ile His Val Ile Leu Thr Ile Glu Ser Ala Ala Met Pro | |
| GAG CGC CCA ATC AAG GTC CGA CGG CTC CCT CGT GAC | 806 |
| Glu Arg Pro Ile Lys Val Glu Leu Leu Pro Arg Arg | |
| TGG TAC TCT TCC CGG AAC CGT GAG TCC CCT GAA ATC TGG | 845 |
| Ile Tyr Ser Ser Pro Asn Arg Glu Phe Pro Glu Ile Trp | |
| GCA CCT GAC CGT CGC GCA ATT GAA GAG CAG AGT GGT | 884 |
| Ala Leu Asp Gly Leu Arg Ala Ile Glu Glu Glu Ser Gly | |

FIG. 2 (1ere planche)

TGG ACC ATT GAG ACC AAC ATT CAO CGG TAC TAC GGC GAT 923
 Trp Thr Ile Glu Asn Ile Glu Glu Tyr Tyr Ala Asp
 AAC AAC GCC ATT GTG CTC CCA ATC GGT GGC GAG AGC 962
 Lys Asn Ala Ile Val Val Leu Pro Ile Gly Glu Ser
 TCC TTC TAC TCT GAC TCG GAA GAG CCA AAC AAC GGC AAG 1001
 Ser Phe Tyr Ser Asp Trp Glu Glu Pro Asn Asn Gly Lys
 AAC TAC CAG TCC GAG ACC TTC CTG ACT CAG GAG CTC GCA 1040
 Asn Tyr Glu Trp Glu Thr Phe Leu Thr Glu Glu Leu Ala
 CGG ATC CTG GAC AAC GGC TTC CGT TCC AAC ACC GAT CGC 1079
 Pro Leu Asp Ser Gly Phe Arg Ser Asn Thr Asp Arg
 GCC ATC ACC GGT ATC TCC ATG GGC GGT ACC GCT GCG GTT 1118
 Ala Ile Thr Gly Ile Ser Met Gly Gly Thr Ala Ala Val
 AAC ATC GCA ACC CAC CAC CCA GAC ATG TTI AAG TTC GTC 1157
 Asn Ile Ala Ile His His Pro Asp Phe Lys Phe Val
 GGT TCC TTC TCC GGC TAT CTG GAC ACC ACC TCC GCT GGC 1196
 Gly Ser Ser Gly Tyr Leu Asp Thr Ser Ala Gly
 ATG CCA ATC GCT ATT TCC GCA GCC CTC GCA GAC GGC GCG 1235
 Met Pro Ile Ala Ile Ser Ala Ala Leu Ala Asp Ala Gly
 GGA TAC GAT GGC AAC GCA ATG TGG GCA GTC GGT TCT 1274
 Gly Tyr Asp Ala Asn Ala Met Trp Gly Pro Val Gly Ser
 GAG CGC TCG CAG GAA AAC GAT CCA AAC ACC AAC GTC GAC 1313
 Glu Arg Trp Glu Asn Asp Pro Lys Ser Asn Val Asp
 AAC CTC AAC GGC ARG ACC ATC TAC GGT TCC TCT GGT AAC 1352
 Lys Leu Lys Gly Thr Ile Tyr Val Ser Ser Gly Asn
 CGT GCA GAT GAC TCC GGT AAG GAA GAC TGT GTC GCT ATT 1391
 Gly Ala Asp Asp Phe Gly Ser Glu Asp Ser Val Ala Ile
 GGA CCT GCA AAC GCG ACA GGT GTC GGT CTG GAA GTT ATC 1430
 Gly Pro Ala Asn Ala Thr Gly Val Gly Leu Glu Val Ile
 TCC CGT ATG ACT TCC CAG ACC TTC GTC GAT CGT GCA AAC 1469
 Ser Arg Met Thr Ser Glu Thr Phe Val Asp Arg Ala Asn
 CAG CCT GGC GTC GAA GTT GTC GCT ACC TTC CGT CCA TCC 1508
 Glu Ala Gly Val Glu Val Val Ala Ser Phe Arg Pro Ser
 CGC GTG CAC TCA TGG GAA TAC TGG CAG TTC GAG ATG ACT 1547
 Gly Val His Ser Trp Glu Tyr Trp Glu Phe Glu Met Thr

FIG.2 (2eme planche)

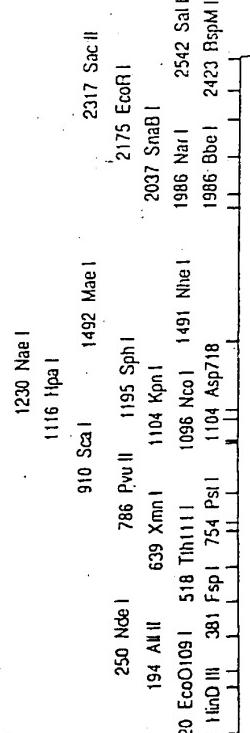
CAG CGG TTC CCT CAC ATC GCT AAC GCT CTC GGC ATG TCC 1586
 Glu Ala Phe Pro His Ile Ala Asn Ala Leu Gly Met Ser
 ACT GAG GAC CGT GGC GTT GAG TGT GCA CCT GTC GGC GCA 1623
 Thr Glu Asp Arg Gly Val Glu Cys Ala Pro Val Gly Ala
 ATC GCT GAC GCT GTT GGC GAC GGC GCG ATG GGC ACC TCC 1664
 Ile Ala Asp Ala Val Ala Asp Gly Ala Met Gly Thr Cys
 CTG ACC AAC GAA TAC GAT GTI ACC GGC GGT AAG GGC CAG 1703
 Leu Thr Asn Glu Tyr Asp Val Thr Gly Gly Lys Ala Glu
 GAC TTC GCT AAC GGT CGG GCA TAC TGG TCT GCA AAC ACT 1742
 Asp Phe Ala Asn Ala Arg Ala Tyr Trp Ser Ala Asn Thr
 GCC GCT TTC GGC CTG GTT GGA CGC ATC AAC GCT CGT TAC 1781
 Gly Ala Phe Gly Leu Val Gly Arg Ile Asn Ala Arg Tyr
 TCT GAG CTG CGT GGA CCT GAC TCC TGG TGT GGC TAC CCA 1820
 Ser Glu Leu Gly Gly Asp Ser Trp Leu Gly Tyr Phe
 ACC TCT TCT GAC TGT AAG ACA CCA GAC CGA CGG CGC CGC 1859
 Thr Ser Ser Glu Leu Lys Thr Pro Asp Gly Arg Gly Arg
 TTC GTC ACC TTC GAG CAC GGC TCC ATC TAC TGA ACC GGC 1898
 Phe Val Thr Phe Glu Gly Ser Ile Tyr Trp Thr Ala
 ACC ACT GGT CCT TGG GAA ATC CCA GGC GAT ATC CTC GGC 1937
 Thr Thr Gly Pro Trp Glu Ile Pro Gly Asp Met Leu Ala
 CGA TGG CGC ACC CAG GAC TAT GAG GGC 'ACC CTC GGC 1976
 Ala Trp Gly Thr Glu Asp Tyr Glu Lys Gly Ser Leu Gly
 TAC CCA ACC GGC GCA GTT GAA TAC AAC GGT GGC CTC 2015
 Tyr Pro Thr Gly Ala Ala Val Glu Tyr Asn Gly Gly Leu
 CGC CAG CAG TTC GAA GGT GGT TAC GTA TTC CGT ACC TCC 2054
 Arg Glu Glu Phe Glu Gly Tyr Val Phe Arg Thr Ser
 AAT AAC CAG TCT TAC TGG GTT CGC GGA GAA ATC TCC AAG 2093
 Asn Asn Glu Ser Tyr Trp Val Arg Gly Glu Ile Ser Lys
 AAG TAC GGC GAT GAC GGA ATC TTC GCT CGG CTC GGT TCC 2132
 Lys Tyr Ala Asp Asp Gly Ile Phe Ala Glu Leu Gly Phe
 CCA ACC GGC AAT CAG AAC TGG ATC AAC GGT CGC GGT TCC 2171
 Pro Thr Gly Asn Glu Lys Leu Ile Asn Gly Gly Ala Phe
 CAG GAA TTC GAA AAC GGC AAC ATC TAC TGG TCC GTC TCC 2210
 Glu Glu Phe Glu Lys Asn Ala Ser Tyr Trp Ser Val Ser

FIG.2 (3eme planche)

ACT GGC GCG CAC GTG ATT CTG CAC GGC GAC ATC TTC GAC 2249
 Thr Gly Ala His Val Ile Leu His Gly Asp Ile Phe Asp
 CGA TGG GGT GCT AAG GGC TGG CAG CGG CGG GAA TAC GCG 2288
 Ala Trp Gly Ala Lys Gly Trp Glu Glu Gly Glu Tyr Gly
 TTC CCA ACC TCT GAC CAG ACC GCA ATC ACC GCG GGT GGA 2327
 Phe Pro Thr Ser Asp Glu Thr Ala Ile Thr Ala Gly Ile
 CAG ACC ATT CAT TTC CAG AAC GGC ACC ATC CGT CAG GTC 2366
 Glu Thr Ile Asp Phe Glu Asn Gly Thr Ile Arg Glu Val
 AAT GGC CCA ATT GAG GAG TCT CCC TAAATGTCGACGCCATCA 2409
 Asn Gly Arg Ile Glu Glu Ser Arg
 CGCAACTCTCGTTCCGGACTTTTGCCCTGACCGCTGCTGCTGGGGGA 2461
 GTCACTGTTGAAAGGAGATGATTCCTCCCTCGACAGGCGCCAGCGAAA 2513
 CGACCGCTGGTCAGCACCAACCGCGAAGGTGCGC

2547

FIG.2 (4eme planche)



CSP1 (HSAl)

FIG. 3

三

The diagram illustrates the construction of the pCGL 613' vector. At the top, a circular plasmid is shown with various restriction sites and their corresponding fragment sizes (in kb) indicated by arrows pointing outwards:

- PstI 0.00
- PstI 7.49
- Hind III 0.46
- KpnI/Clal 1.56
- Ap csp1
- pCGL 613'
- 8.25 Kb
- Tet
- NruI 5.92
- Km
- BamHI 5.32
- KpnI/Clal 3.06
- EcoRI/SmaI 4.56
- EcoRI 4.12

A scale bar indicates 1 kb. Below the plasmid, a horizontal line represents the linearized vector. The restriction sites and their positions are labeled as follows:

- wt: HindIII, Ncol, EcoRI, KpnI, SmaI, HindIII, EcoRI, BamHI.
- BgIII: BgIII (at both ends), 7.5 kb (between HindIII and SmaI).

An arrow points from the linearized vector down to a second horizontal line representing the csp1 construct. This construct has the following features:

- csp1: HindIII, Ncol, Km, EcoRI, KpnI, EcoRI.
- BamHI: SmaI, HindIII, BamHI.
- 4.5 kb: A scale bar indicating the size of the csp1 construct.

FIG. 5

PERSCRIBSH

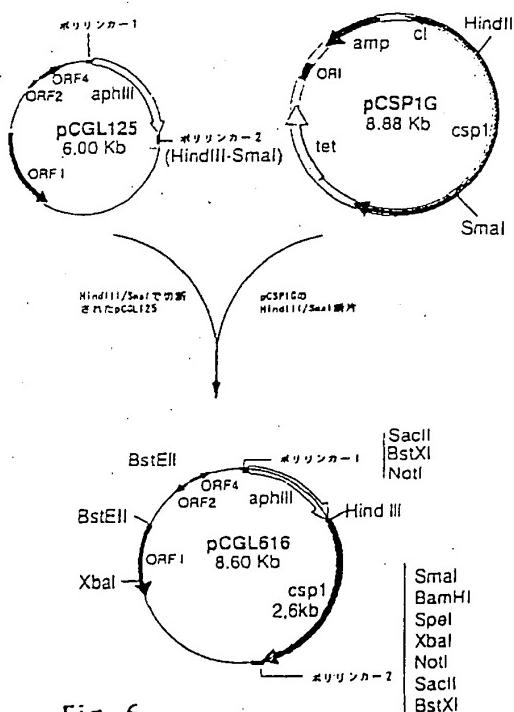


Fig. 6

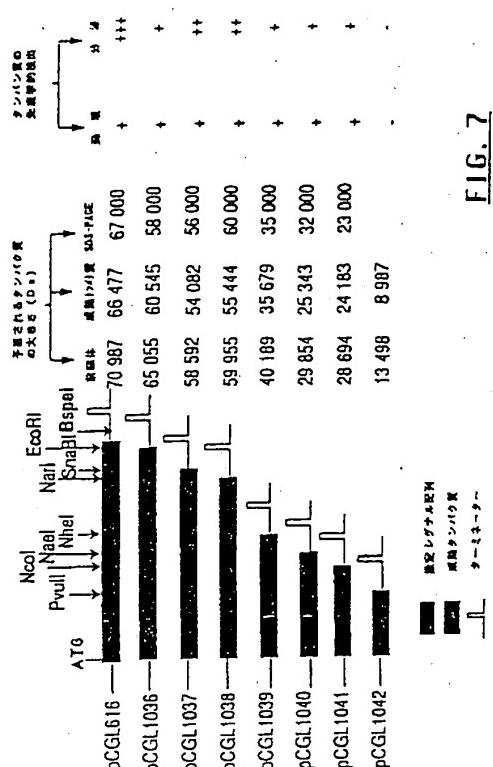


Fig. 7

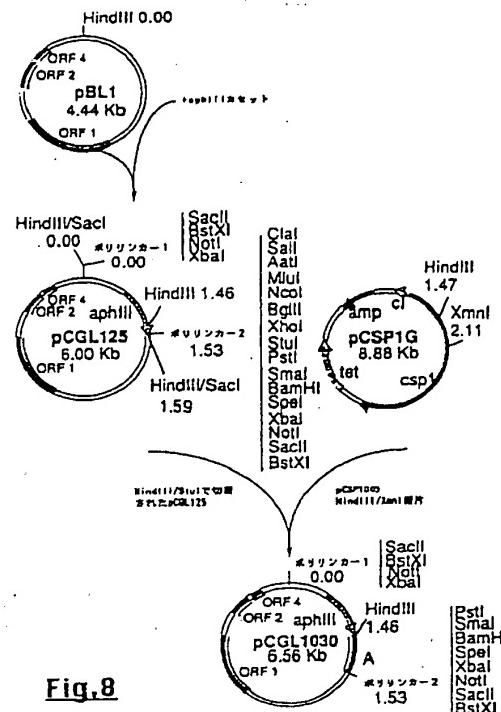


Fig. 8

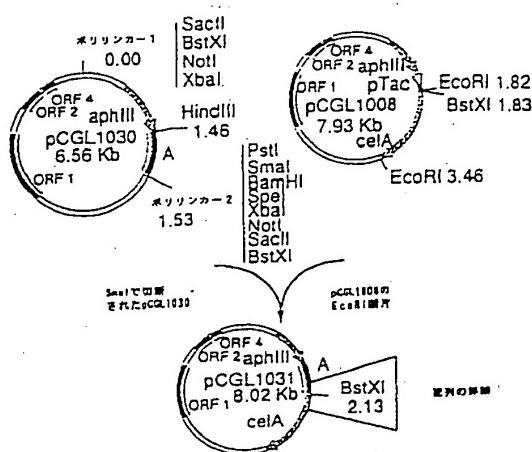


Fig. 9

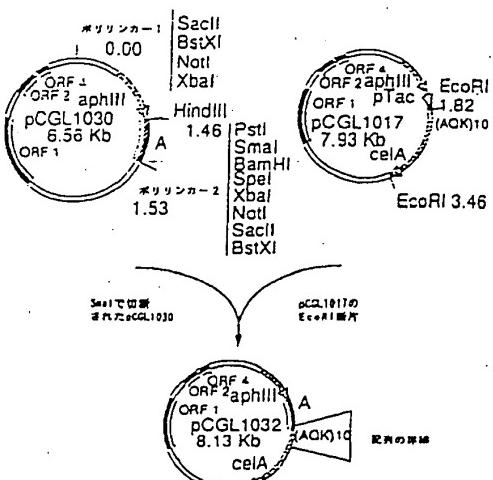


Fig. 10

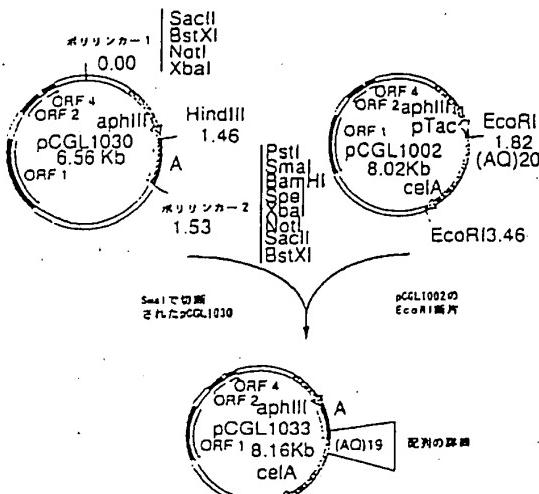


FIG. 11

ATC ACC AAG ACC CGC GAG TCC GTT GCT TAC GCA CTC AAC 1001
 Ile Thr Lys Thr Arg Glu Ser Val Ala Tyr Ala Leu Lys

 GTT GAC CAG GAA GCT ACC CCT GCT TTC CAG GCA TAC CCC 1040
 Val Asp Glu Ala Thr Ala Ala Phe Glu Ala Tyr Arg

 AAC GCA CCT CGC GAT GCA GCT ATC TCT ATC AAC CCA GAT 1079
 Asn Ala Leu Arg Asp Ala Ala Ile Ser Ile Asn Pro Asp

 CGC TCT ATC AAC CCA GAT ACC TCT ATC AAC CTA CTG ATC 1118
 Gly Ser Ile Asn Pro Asp Thr Ser Ile Asn Leu Ile

 CAT CCT GCT AAC GCT AAC ACC ACC GAT CGT GCA GAG 1157
 Asp Ala Ala Asn Ala Ala Asn Arg Thr Asp Arg Ala Glu

 ATC GAG CAT TAC GCT CAC CCT TAC ACC CAG ACC GAT ATT 1196
 Ile Glu Asp Tyr Ala His Leu Tyr Thr Glu Thr Asp Ile

 GCT CCT GAA ACT CCA CAG CCT GCA TAC CCT TTC CAG GAC 1235
 Ala Leu Glu Thr Pro Glu Leu Ala Tyr Ala Phe Glu Asp

 CTG AAG GCT CCT CAG GCT GAG GTC GAC GCA GAC TTC GAG 1274
 Leu Lys Ala Leu Ala Glu Val Asp Ala Asp Phe Glu

 TGG TTG CGC GAG TTC GGA ATC GAC CAG GAA GAC GGT AAC 1313
 Tyr Leu Glu Phe Glu Ile Asp Glu Glu Asp Glu Asn

 TAC CCT CAG CGC TAC CAC CCT CCT GCT GTC GAG GCA CTC 1352
 Tyr Val Glu Arg Tyr His Leu Pro Ala Val Glu Ala Leu

 AAG CCT GAG GTC GAC GCT CGC GTC GCA GCA ATT GAG CCA 1391
 Lys Ala Glu Val Asp Ala Arg Val Ala Ile Glu Pro

 CCT CGT GCA GAC TCC ATC GCT AAC CCT GAG GCG CAG 1430
 Leu Arg Ala Asp Ser Ile Ala Lys Asn Leu Glu Ala Glu

 AAG CCT GAC GTC CTG GTT CGC CAG CTC TTC CTC GAG CGT 1469
 Lys Ser Asp Val Leu Val Arg Glu Leu Phe Leu Glu Arg

 GCA ACC GCA CAG CGC GAC ACC CTC CCT GTT GTC GAG GCG 1508
 Ala Thr Ala Glu Arg Asp Thr Leu Arg Val Glu Ala

 ATC TTC TCT ACC TCT GCT CGT TAC CCT GAA CTC TAC GAG 1547
 Ile Phe Ser Thr Ser Ala Arg Tyr Val Glu Leu Tyr Glu

 AAC GTC GAG AAC GTT AAC GTT GAG AAC AAG ACC CCT CGC 1586
 Asn Val Glu Asn Val Asn Val Glu Asn Lys Thr Leu Arg

FIG. 12 (2eme planche)

GAATTCCCTGTAATTAACCGCGTTTACTACTTTTCAGGGGTCTATCTTAC 52
 CAGATCGTCAGTTGGGTAGACTCACCTGAATAATAATTGCAACCGAACGG 104
 GTGATATATGCCCTAATTGGCTCAAGTAGCTTGAGGTAAAGGTATTTTGGTG 156
 AACAAATTTCAGCTCCGGTAAAGACTTTCTATGCCCTCAGACCTCTAT 208
 TAGGAAATCTGACACCACCTTGAATTAATAGCTTACCCCGAATGGGGGAG 260
 GGTCAATTGCTGTAAGGTAGTTTCTGATGCAATATGACCTGGTTATAA 312
 AGAAATGTTAAACGTCAGATCGATATAAAAAGAACACTTGTACTTCAGT 364
 TTGAGGCAATTCTCCGATTGGCTGGCAAAAAATCTCAATTGTGCTTACAG 416
 TTTCTCIAACGACAGGGCTGGTAAGCTGCTACTTGGTGGCTTAGTGACTGG 468
 CGTTTACTGGTATAAAGTAATCCCATGTCGATCAGGCAATTGGGTTCT 520
 TCCATAGGCAATCCAAAGGTTTGTCTTTCGATACCTTACAAGGAGCTTC 572
 GCCTCT ATG TTT AAC AAC CGT ATC CGC ACT GCA CCT CTT 611
 Met Phe Asn Asn Arg Ile Arg Thr Ala Ala Leu

 GCT GCT GCA ATC OCA ATC TCC AAC GCA GCT TCC CGC CCT 650
 Ala Ile Ala Ile Ala Ser Thr Ala Ala Ser Glu Val

 GCT ATC CCT GCA TTC OCT CAG CAG ACC AAC CCA ACT TTC 689
 Ala Ile Pro Ala Phe Ala Glu Glu Thr Asn Pro Thr Phe

 AAC ATC ACC AAC GGC TTC AAC GAT GCT GAT GCA TCC ACC 728
 Asn Thr Asn Gly Phe Asn Asp Ala Asp Gly Ser Thr

 ATC CGC CCT GGC CCT GTT AAC CAC ACC GAG GAA ACC 767
 Ile Glu Pro Val Glu Pro Val Asp His Thr Glu Glu Thr

 CTC CGC GAC CGT ACT GAG TCC ACC GCG CCT TAC CCT GAA 806
 Leu Arg Asp Leu Thr Asp Ser Thr Gly Ala Tyr Leu Glu

 GAG TTC CAG AAC GGC CCT GAG GAA ATC CCT GAA GCA 845
 Glu Phe Glu Asn Glu Asn Val Glu Glu Ile Val Glu Ala

 TAC CGC CCT GAG CCT TCC GCA GAC GCA TTC GAT CCT 884
 Tyr Leu Glu Val Glu Ala Ser Ala Asp Glu Phe Asp Pro

 TCT GAG CAG CCT TAC GAG CCT TTC GAG CCT CCT CCT GGC 923
 Ser Glu Glu Ala Ala Tyr Ala Phe Glu Glu Ala Ala Arg

 GTC CCT GCA TCC CAG CAG CCT GCA CCT CCT CCT GAG ACC 962
 Val Arg Ala Ser Glu Glu Leu Ala Ala Ser Ala Glu Thr

FIG. 12 (3eme planche)

CAG CAC TAC TCT GCG CTG ATC CCT AAC CCT TTC ATC GCA 1625
 Glu His Tyr Ser Ala Leu Ile Pro Asn Leu Phe Ile Ala

 GCA GTT GCA AAC ATC ACC GAG CTC AAC CCT GCA GAT CCT 1664
 Ala Val Ala Asn Ile Ser Glu Leu Asn Ala Ala Asp Ala

 GCA GCA GCA CCT TAC TAC CTC CAC TGG GAC ACC GAC CCT 1703
 Glu Ala Ala Ala Tyr Tyr Leu His Trp Asp Thr Asp Leu

 GCA ACC AAC GAT GAG GAC GAA GCT TAC TAC AAC CCT AAC 1742
 Ala Thr Asn Asp Glu Asp Glu Ala Tyr Tyr Lys Ala Lys

 CTC GAC TTC GCT ATC GAG ACC TAC GCA AAG ATC CTG TTC 1781
 Leu Asp Phe Ala Ile Glu Thr Tyr Ala Lys Ile Leu Phe

 AAC CCT GAA GTT TGG CAG CCT GCA CCT CCT CCT AAC CCT CAG 1820
 Asn Glu Val Trp Glu Glu Pro Leu Ala Tyr Val Glu

 AAC CTG GAT GCA GGC GCA CCT CAG GAA GCA CCT AAC CCT 1859
 Asn Leu Asp Ala Gly Ala Arg Glu Ala Ala Asp Arg

 GAG CCT CCT GCA CCT GAC GAA CCT TAC CCC CCT CCT GAG 1898
 Glu Ala Ala Arg Ala Ala Asp Glu Ala Tyr Arg Ala Glu

 CAG CTC CGC ATC GCT CAG GAA GCA CCT GAC CCT CAG AAC 1937
 Glu Leu Arg Ile Ala Glu Glu Ala Ala Asp Ala Glu Lys

 GCT ATC CCT GAG GCG CCT CCT GAG GAA GCA GAC GGC CCT 1976
 Ala Ile Ala Glu Ala Leu Ala Lys Glu Glu Ala Glu Lys

 AAC GAC AAC TCC CAC GAC AAC ACC GAG ACC CCT TCT TCT 2015
 Asn Asp Asn Ser Ser Asp Asn Thr Glu Thr Gly Ser Ser

 GAC ATC GCA TCC TGG GCA CCT TTC GCA GCA ATT GCA CCT 2054
 Asp Ile Glu Ser Trp Glu Pro Phe Ala Ala Ile Ala Ala

 ATC ATC GCA ATC GCA ATC GCT ATC TTC GCA TTC CTC TCC 2093
 Ile Ile Ala Ala Ile Ala Ile Ile Phe Pro Phe Leu Ser

 GGT ATC GTT AAG TTC TAA TTTCGAAACGGAGATAGCTAAAGTTAAA 2139
 Glu Ile Val Lys Phe

 GAAAGCCACCTCCCTGATCTAACGGAGGTGGCTTACAGGACCAAAA 2191

 GAGCCGGAGGTTGGCTCGATTAAGCAAAAAATCTTTGCTTAAAGGAAACGT 2243

 GAGCCGGAGGTTGGCTCGATTAAGCAAAAAATCTTTGCTTAAAGGAAACGT 2295

FIG. 12 (3eme planche)

FIG. 12 (4eme planche)

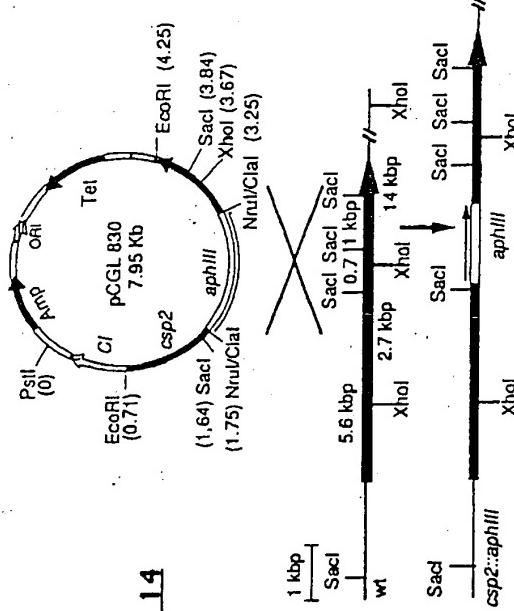


Figure 14

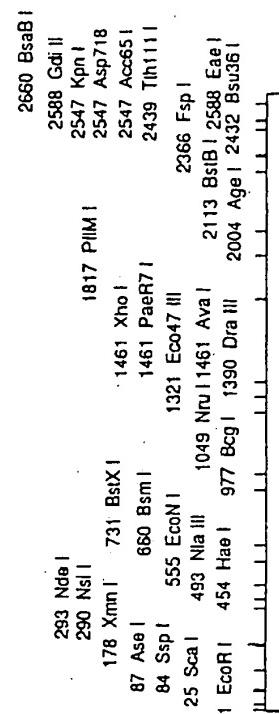


FIG. 13

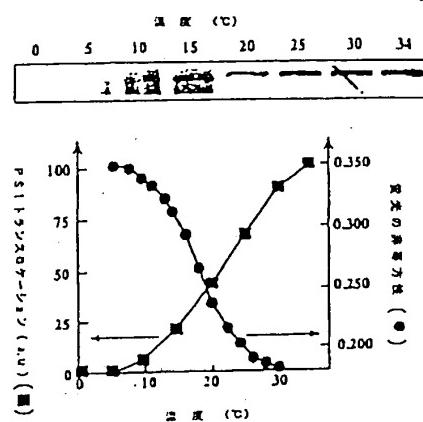


FIG. 15

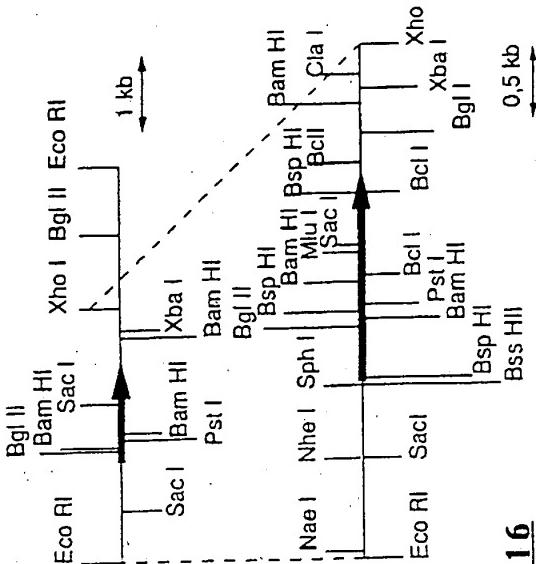


FIG. 16

TTC GAC TTC GAC CCT AAC GCC AAC TCC CAT CTG GAA ATC 1001
 ser asp phe asp pro lys gly lys ser asp leu glu ile

 ATG CGT TTC TGC CAG TCC TTC ATG ACC CAG CTG CAC CCC 1040
 met arg phe cys gln ser phe met the glu leu his arg

 CAC ATC GGT GAG TAC CGC GAG GTT CCT GCA GGT GAC ATC 1079
 his ile gly gln tyr arg asp val pro ala gly asp ile

 GGA GTT CGT GCG CGC GAG ATC GGT TAC CTG TTT GGC CCC 1118
 phe val gly gln arg gln ile gly tyr leu phe gln his

 TAC CCT CCC ATG GCC AAC CAG CAC GAG TCC GGC GTC TTG 1157
 tyr arg arg met ala asn gln his glu ser gly val leu

 ACC GGT AAC GGC CTG ACC TGG CCT GCA TCC CTG CCC 1196
 the gly lys leu thr trp gly gln ser ile val asp arg

 ACC GAG GCA ACT GGC TAC GGC TCC GTT TAC TTC GTG ACT 1235
 the glu ala the gly tyr gly cys val tyr phe val ser

 GAA ATG ATC AAC GCT AAC GGC GAG AGC ATC ACC GGC CAG 1274
 gln met ile lys ala gln glu ser ile ser gly gln

 AAC ATC ATC GTC TCC GGT TCC GGC AAC GTC GCA ACC TAC 1313
 lys ile ile val ser gly ser gln asp val ala the tyr

 GCG ATT GAA AAC GGT CAG GAA CTC GGC GCA ACC GTT ATC 1352
 ala ile gln lys ala gln gln leu gly ala thr val ile

 GGT TTC TCC GAT TCC AGC GGT TGG GTT CAT ACC CCT ATT 1391
 gly phe ser asp ser ser gln trp val his pro arg

 GGC GTT GAC GTG GCT AAC CCT COC GLA ATC AAC GAA GTT 1430
 gly val asp val ala lys leu arg gln lle lys glu val

 CCC CCC GCA CGC GTA TCC GTG TAC GGC GCA GAC GTT GAA 1469
 arg arg ala arg val ser val tyr ala asp glu val glut

 GGC GCA ACC TAC GAC ACC GAC CGG TCC ATC TGG GAA CTC 1508
 gly ala the tyr his thr asp gly ser ile trp asp leu

 AAG TGC GAT ATC GCT CTT CCT TGT GCA ACT CAG AAC GAG 1547
 lys cys asp ile ala leu pro tyros ala the gln asn glu

 CTC ATC GGT GAG AAC GCT AAC ACT CTT GCA GAC AAC GGC 1586
 leu asn gln gly asn ala lys the leu ala asp asn gly

 TGG CGT TTC GTT GCT GAA CGC GCG AAC ATC ATG CCT TCC ACC 1623
 cys arg phe val ala gln gly ala asn met pro ser thr

FIG. 17 (2^eme planche)

GCTAGCCCTGGGACTCTAGAGATTGTCAAAACGGCTCAAATTCTCGGA 52
TGCAGGCCATAAAGAGTCGACCAATTCCATTTGAGGGTGTCAAGTGTGG 104
CCAGGTTATAAACCGACTGCTCACCTGGTCTCATTCGGTGTGCGATGAA 156
TTAATTAAAAGAAGAGACTTCATGCMGTACCCCGGTTCGGCATACACM 208
TTGATTAACCTAAAGAAATTTCACAACTATTAACTCTTGTGGTCATAC 260
TGTGCGACACTGCGATTAATGACAGTGACCGATTACCCGGTAAATGCCCG 312
AGTGAGTTAACCTTCAGCAGAACGTTCTTCAGGGCATCCGTAGTTA 364
AAACTTAAACCGTGGATGACAGRCGGGTGAGGTCAACCSAGTTTA 416
AAAGTTTCACCGATACATTTTCACCGCAATTGGTCAGCAGAACCCCTAG 468
GATGCTACATGCGCCCTCAATGGCAACCCAAACATCACTAAATGGCCAGA 520
TACACCTTTAAATTCGTGGCCGATCAGCCGAGATGGCAATAGGAAATC 572
AIG ACA GTT GAT GAG CAG GTC TCT AAC TAT TAC GAC ATG 611
met the val asp glu val ser asp tyr asp arg
CTT CTG AAG CGC AAT GCT CGC GAG CCT GAA TTT CAC CAG 630
leu leu lys arg asn ala gly glu pro phe his gln
GCA GTG GCA GAG GTT TTC GAA TCT TTT AAG ATC GTC CTG 669
ala val ala glu val leu glu ser leu lys ile val leu
GAA AAG GAC CCT CAI ZAC GCT GAT TAC GOT CGC ATC CAG 728
glu lys asp pro his tyr ala asp tyr gly leu leu leu gln
CCC CTG TCC GAG CCT GAG COT CGA CTC ATC TTC COT CTG 767
arg leu cys glu pro glu arg gln leu ile phe arg val
CCT TCG GTZ GAT GAC CAG GCG GAC GTC GAC GTC AAC CCT 804
pro trp val asp asp glu gln gln val his val asn gln
GGT STC CGC GTG CAG TTC AAC TCT GCA CTT GGA CCT TAC 845
gly phe arg val gln phe asn ser ala leu gln pro tyr
AAG GCC GGC CTG CGC TTC AAC CCA TCT GTA AAC CTG GGC 884
lys gly gln arg phe his pro ser val asn leu gln asn
ATT GTG AAG TTC CTG CGC TTT GAG CAG ATC TTT AAA AAC 923
ile val lys phe leu gln phe glu gln ile phe lys asn
TCC TCA ACC GGC CCC TCC CCA ATC GCT GGT GCG AAC AGG CCT GGT GCA 962
ser leu the gly leu pro ilo ala gly gly gln lys gln gly

FIG. 17 (1ere planche)

CCA GAG GCT GTT CAG GTC TTC CGT GAG CCC GAC ATC CCC 1664
 pro glu ala val glu val phe arg glu arg asp ile arg

 TTC CGA CGA CGG AAC GCA CCT AAC CCT GGT GGC TGA GCA 1703
 phe gly pro gly lys ala ala asp ala gly gly val ala.

 ACC TCC GCT CTG GAG ATG CAG CAG AAC CCT TCG CCC GAT 1742
 thr ser ala leu glu met gln gln asn ala ser arg asp

 TCC TGG AGC TTC GAG TAC ACC GAC GAG CCC CTC GAG GTG 1781
 ser trp ser phe glu tyr thr asp glu arg leu gln val

 ATC ATG AAG AAC ATC TCT AAC ACC TGT GCA GAG ACC GCA 1820
 ile met lys asn ile phe lys thr cys ala glu thr ala

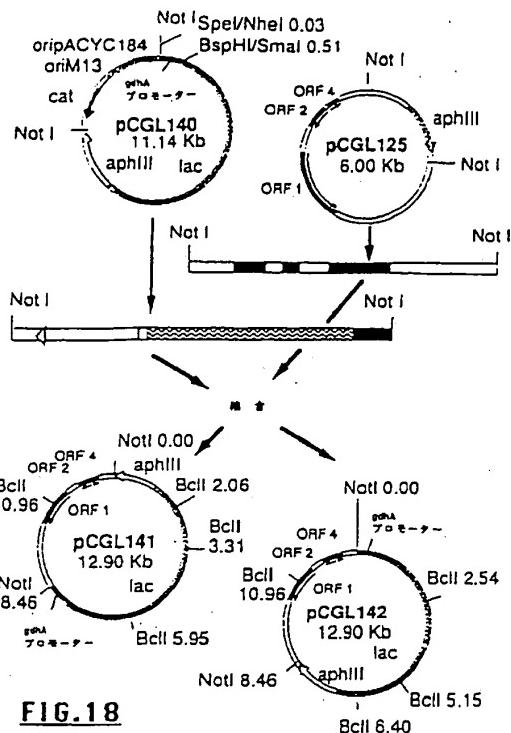
 GCA GAG TAT GGA CAC GAG AAC CAT TAC GTT GTC GGC GCT 1859
 ala glu tyr gly his glu asn asp tyr val val gly ala

 AAC ATT GCT GCC TTC XAG AAC GTA CCT GAC GCG AGC CTG 1898
 asn ile ala gly phe lys lys val ala asp ala met leu

 GCA GAG CGC CTC ATC TAA GACCCCTGCACTTAACTTAAACCCCTGA 1944
 ala gln gly val lle och

 TCCGGCTTAAGGATCAGGGATTTTTGATTTCTTCAGGTCATTAACTCCGATC 1996
 CACATGGGTTAACGACGGTGTGGCGTGGCCGATGATGATCACCCGTGTCT 2048
 TTAAAGCTGGCCAGAGCTGGGAAAGATCCGGTCAATGAGGCCCATCTTGT 2100
 GGCTGGCTGGCTTCATCGACAAATCACTACCTGAGGGGTGGGTCGCCAAAGCAGC 2152
 CCCTTGGCTGGCTTCATCGACAAATCACTACCTGAGGGGTGGGTCGCCAAAGCAGC 2190

FIG. 17 (3eme planche)



DGF1 : 32 mer
 $5' \text{AATCCATGGCAATGGCGGCCCTGTCGACCCC} 3'$

DGF2 : 28 mer
 $5' \text{GGGGTCCACAGGCCGGGATTGCCATGG} 3'$

DGF3 : 120 mer
 $5' \text{CAGGCACAGGCTCAGGCCAGGGCACAGGCCAGGGCAGG}$
 $\text{CCCCAGGCCAGGCAGGCCACAGGCCAGGCCAGGCCAGG}$
 $\text{GGCACAGGCCACAGGCTCAGGCCAGGCCAGGCCAGGCCA}$

DGF6 : 120 mer
 $5' \text{TGAACCTTAGGCTGC3CCT3AGGCTTGTCCTGTCCTGCG}$
 $\text{CTGCGGCTGGCGCTGCGCTGAGGCTGGGCGCTGGCGCTG}$
 $\text{CGCGCTGGGCGCTGTGCGCTGGCGCTGAGGCTGTGCG}$

FIG.19

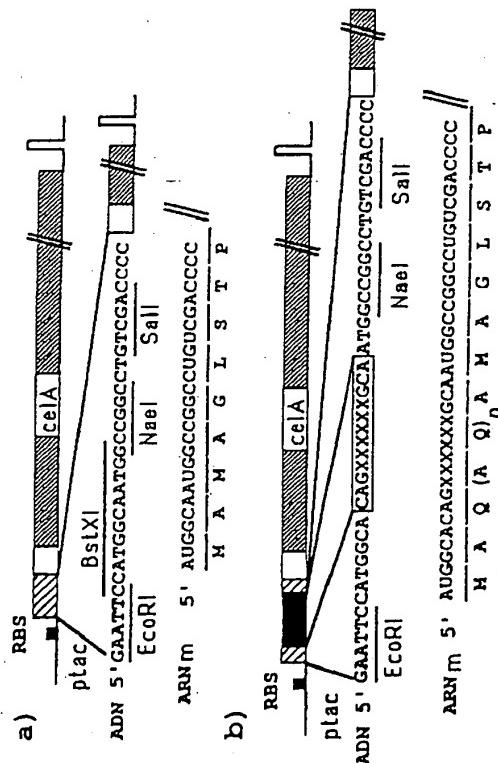
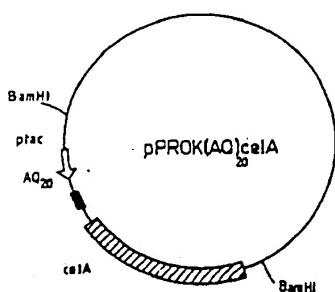
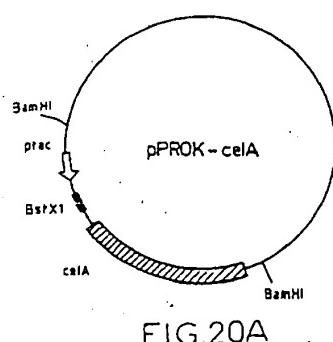


FIG.21

手 感 情 正 著

平成 5 年 8 月 22 日

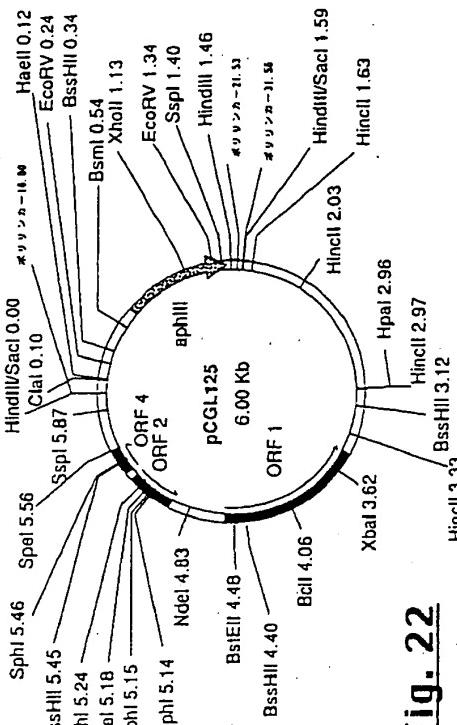


Fig. 22

| | | | |
|---|---|--|--|
| A. CLASSIFICATION OF SUBJECT MATTER | | International application No. PCT/FR 92/00744 | |
| Int.Cl. 5 | C12M15/77; C07K13/00; | C12N1/21; C12N15/90 A61K39/50 | |
| According to International Patent Classification (IPC) or to both national classifications and IPC | | | |
| B.田字格 SEARCHED | | | |
| Inventor's name does not appear under classification system as defined by classification systems | | | |
| Int.Cl. 5 | C12H; | C07K | |
| Document number which can be used to determine whether such documents are included in the file | | | |
| Information data is extracted during an automatic search (name of document and, where possible, source from which it was obtained) | | | |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | | | |
| Category | Classification of document, with indication, where appropriate, of the relevant pages | Relevant to patent No. | |
| A | WO A 8 E09 821 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY) 15 December 1988 | 1-34 | |
| A | EP A 2 372 492 (ASAMI KASEI KOGYO KABUSHIKI KAISHA) 4 July 1986 | 1-34 | |
| A | BIO/TECHNOLOGY vol. 8, No. 6, June 1990, NATURE AMERICA, INC., NEW YORK, US pages 359 - 363 R.-A. PETIT ET AL.: "Hypersecretion of a cellulase (endoglucanase) by thermophilic <i>Escherichia coli</i> by induction of chromosome DNA amplification" | 1-34 | |
| | | -/-/- | |
| <input type="checkbox"/> | Further documents are listed in the continuation of Part C. | <input type="checkbox"/> See patent family area. | |
| <p>Part C contains a list of documents which, in the opinion of the Inventor, are relevant to his/her invention. The list is not exhaustive. It is not intended to give an indication of the importance of each document. The priority date underlying the International Application is not mentioned in Part C.</p> <p>The following information is given for each document:</p> <ul style="list-style-type: none"> -¹ document number and title; -² date of publication; -³ name of the inventor(s); -⁴ name of the assignee(s); -⁵ date of recording of the document in the patent office or other document office; -⁶ document published prior to the International filing date but later than the priority date. | | | |
| Date of the earliest completion of the International search | | Date of sending of the International search report | |
| 19 October 1992 (39.10.92) | | 4 November 1992 (04.11.92) | |
| Place and mailing address of the EA/European Patent Office Postnach No. | | Assignment officer Telephone No. | |
| Form PCT/FR/00744 (second sheet of four sheets) | | | |

| C (Continuation): DOCUMENTS CONSIDERED TO BE RELEVANT | | International application No. PCT/EP 92/00744 |
|---|--|--|
| Category | Character of document, with indication, where appropriate, of the relevant passage | Reference to claim No. |
| A | BIO TECHNOLOGY vol. 346, page 1 January 1991, Nature America, INC., NEW YORK; US pages 84 - 87 A. SCHWARZER AND A. PUNLER "Manipulation of <i>Corynebacterium glutamicum</i> by gene disruption and replacement" ----- | 1-34 |

国際特許報告

| | | | | |
|---------------------------------|------------------|---------------|------------------|----------|
| Patent document number reported | Publication date | Patent holder | Publication date | |
| WO-A-8003821 | 16-12-88 | US-A- | 4965197 | 23-10-90 |
| FR-A-1575492 | 04-07-88 | JP-A- | 61268105 | 27-11-86 |

For more details about this report, see Official Journal of the European Patent Office, No. 15/82

フロントページの続き

| | | | |
|---------------------------|---|----------|--|
| (51) Int.Cl. ⁵ | 識別記号 | 序内整理番号 | F I |
| C 12 N 15/90 | | | |
| C 12 P 21/00 | 8214-4B | | |
| 21/08 | 8214-4B | | |
| //(C 12 N 1/21 | | | |
| C 12 R 1:13) | | | |
| (C 12 N 1/21 | | | |
| C 12 R 1:15) | | | |
| (C 12 P 21/00 | | | |
| C 12 R 1:13) | | | |
| (C 12 P 21/00 | | | |
| C 12 R 1:15) | | | |
| (81) 指定国 | EP(AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, N L, SE), JP, US | (72) 発明者 | デュシロン, フランシス フランス國アポン、リュ、メルモ、25、レ ジダンス、ラ、フォンテーヌ、オ、ボワ |
| (72) 発明者 | レブロン, ジエラール フランス國レ、ジュリ、リュ、デ、バー ト、5 | (72) 発明者 | ルノー, ミシェル フランス國レ、ジュリ、リュ、デ、コー ス、23 |